MONAURAL AND BINAURAL DETECTION OF SPEECH AND ITS RELATION TO THE SPEECH MODE HYPOTHESIS

By
WILLIAM G. THOMAS

A DISSERTATION PRESENTED TO THE GRADUATE COUNCIL OF
THE UNIVERSITY OF FLORIDA
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA
1968

TO JOY LYNNE, WENDY, BILL,
AND KIMI

ACKNOWLEDGMENTS

The author expresses deep gratitude to his advisor, chairman, and friend, Dr. John F. Brandt, for his continued support and guidance, not only throughout the planning and execution of this investigation, but, also, throughout the author's academic career at the University of Florida. The author is further indebted to the members of the dissertation committee for their interest, constructive criticism, and special contributions to the ideas expressed herein.

The author further wishes to acknowledge the invaluable support of the faculty, staff, and students of the Communication Sciences Laboratory of the University of Florida, with a special note of gratitude to Dr. Donald Teas, who read and criticized the original manuscript and made many valuable suggestions.

The author is further indebted to Dr. Robert Scholes, who synthesized the original consonant-vowel tapes while at Haskins Laboratory and permitted their use in this study.

The author also expresses his appreciation and gratitude to his wife and family, who, although enduring many hardships in this pursuit, provided steady encouragement, inspiration, and support.

Finally, the author expresses his appreciation to the Vocational Rehabilitation Administration for financial support in the form of a Special Rehabilitation Research Fellowship. In addition, the research was supported from funds provided by the Graduate School of the University of Florida and the National Institutes of Health (Grant NB 06459).

TABLE OF CONTENTS

		Page
LIST OF TABLES		vi
LIST OF FIGURES		x
CHAPTER		
I	INTRODUCTION	1
II	PROCEDURE	11
	Subjects	11
	Stimuli	11
	Test Stimuli	13
	Experimental Apparatus	14
	Experimental Session	14
	Confidence Ratings	16
III	RESULTS	19
	Monaural Condition	20
	Ear Effect	20
	Effects of Stimulus Types	20
	Effects of Probability for the Occurrence of Speech	23
	Stimulus Omissions	23
	Response Errors	23
	Summary of Results in Monaural Conditions	26
	Binaural Conditions with Directions	26
	Ear Effect	26
	Effects of Stimulus Types	26
	Effects of Probability for the Occurrence of Speech	29
	Stimulus Omissions	29
	Response Errors	30
	Summary of Results in Binaural Conditions with Directions	32
	Binaural Conditions without Directions	32
	Ear Effect	32
	Effects of Stimulus Types	33
	Effects of Probability for the Occurrence of Speech	33
	Response Errors	36
	Summary of Results in Binaural Conditions without Directions	36

Page

Confidence Ratings 38
Monaural Conditions 38
Binaural Conditions with Directions 41
Summary of Results of Confidence Ratings 46

IV DISCUSSION 49
Speech Versus Non-Speech 49
Ear Effect 50
Effects of Stimulus Types 53
Confidence Ratings 55

V SUMMARY 57

APPENDIX A DESCRIPTION OF THE CV, V, VN, AND NBN STIMULI 61

APPENDIX B GENERATION OF TEST TAPES 68

APPENDIX C STATISTICAL SUMMARY TABLES 82

BIBLIOGRAPHY 96

BIOGRAPHICAL SKETCH 101

LIST OF TABLES

Table		Page
1.	Percent correct responses for right and left ears in the monaural condition for each of the three a priori probabilities.	22
2.	Average number of omissions for each stimulus type during an experimental session of 100 stimuli. The omissions are averaged over the three a priori probabilities.	24
3.	Average number of errors during an experimental session for the three a priori probabilities of speech occurrence and the different experimental conditions.	25
4.	Percent correct responses for right and left ears in binaural condition with direction for each of three a priori probabilities.	28
5.	Response errors in the reporting ear (in percent) when competing stimuli are presented in the contralateral ear.	31
6.	Percent correct responses for right and left ears in the binaural condition without direction for each of three a priori probabilities.	35
7.	Response errors in the binaural without direction condition (in percent).	37
8.	Optimum values computed for the monaural right and monaural left conditions.	39
9.	Mean optimum values and standard deviations for four stimulus types in the monaural right and monaural left conditions. The effects of a priori probability of speech have been averaged.	40
10.	Optimum values computed for the binaural with direction condition	44
11.	Mean optimum values and standard deviations for the four stimulus types in the binaural with direction condition. The effects of a priori probability of speech have been averaged.	45

Table Page

12. Total number of the four stimulus types appearing on channel A and channel B of Test Tape I [P(Sp) = 0.25], and total number of times the four stimulus types are in competition with each other in the dichotic condition. 74

13. Total number of the four stimulus types appearing on channel A and channel B of Test Tape II [P(Sp) = 0.25], and total number of times the four stimulus types are in competition with each other in the dichotic condition. 75

14. Total number of the four stimulus types appearing on channel A and channel B of Test Tape III [P(Sp) = 0.50], and total number of times the four stimulus types are in competition with each other in the dichotic condition. 76

15. Total number of the four stimulus types appearing on channel A and channel B of Test Tape IV [P(Sp) = 0.50], and total number of times the four stimulus types are in competition with each other in the dichotic condition. 77

16. Total number of the four stimulus types appearing on channel A and channel B of Test Tape V [P(Sp) = 0.75], and total number of times the four stimulus types are in competition with each other in the dichotic condition. 78

17. Total number of the four stimulus types appearing on channel A and channel B of Test Tape VI [P(Sp) = 0.75], and total number of times the four stimulus types are in competition with each other in the dichotic condition. 79

18. Total number of the four stimulus types appearing on channel A and channel B of all Test Tapes combined, and total number of times the four stimulus types are in competition with each other in the dichotic condition. 80

19. Average relative intensity for channel A and channel B of each test tape. 81

20. Mean correct score and its standard deviation for the right and left ears for the monaural condition with [P(Sp) = 0.25]. The proportion scores represent arcsin transformations. 83

Table		Page
21.	Mean correct score and its standard deviation for the right and left ears for the monaural condition with [P(Sp) = 0.50]. Scores represent arcsin transformations.	84
22.	Mean correct score and its standard deviation for right and left ears in the monaural condition with [P(Sp) = 0.75]. Scores represent arcsin transformations.	85
23.	Summary table for the analysis of variance of the four stimulus types in the monaural conditions.	86
24.	Summary table of Scheffé's test for a posteriori comparisons between means for monaural conditions.	86
25.	Summary table of the analysis of variance for the three a priori probabilities in the monaural conditions.	87
26.	Mean correct score and its standard deviation for binaural right and binaural left conditions with [P(Sp) = 0.25]. Scores represent arcsin transformations.	87
27.	Mean correct score and its standard deviation for binaural right and binaural left conditions with [P(Sp) = 0.50]. Scores represent arcsin transformations.	88
28.	Mean correct score and its standard deviation for binaural right and binaural left conditions with [P(Sp) = 0.75]. Scores represent arcsin transformations.	89
29.	Summary table of the analysis of variance of the four stimulus types in the binaural condition with directions.	90
30.	Summary table of Scheffé's test for a posteriori comparisons between means in the binaural conditions with directions.	90
31.	Summary table of the analysis of variance of the three a priori probabilities in the binaural condition with directions.	91
32.	Mean correct score and its standard deviations for binaural condition without directions with [P(Sp) = 0.50]. Scores represent arcsin transformations.	91

Table		Page
33.	Mean correct score and its standard deviations for binaural condition without directions with [P(Sp) = 0.75]. Scores represent arcsin transformations.	92
34.	Mean correct score and its standard deviations for binaural condition without directions with [P(Sp) = 0.25]. Scores represent arcsin transformations.	93
35.	Summary table of the analysis of variance of the four stimulus types in the binaural condition without directions.	93
36.	Summary table of Scheffé's test for a posteriori comparisons between means in binaural condition without directions.	94
37.	Summary table of analysis of variance for the three a priori probabilities in the binaural condition without directions.	94
38.	Summary table of Scheffé's test for a posteriori comparisons between means in the binaural condition without directions.	95

LIST OF FIGURES

Figure		Page
1.	Schematic diagram of equipment used in each experimental session.	15
2.	Mean percent correct responses for right and left ears in the monaural listening conditions for the four stimulus types. The values are averaged over a priori probabilities.	21
3.	Mean percent correct responses in right and left ears for the four stimulus types in the binaural conditions with direction. The values are averaged over a priori probabilities.	27
4.	Mean percent correct responses in right and left ears for the four stimulus types in the binaural conditions without direction. The values are averaged over a priori probabilities.	34
5.	Probability of a given confidence rating, $P(R_j)$, in the monaural conditions as a function of scale values, R_j. The parameter is stimulus type. The effects of ears and a priori probabilities have been averaged.	42
6.	Type-II ROC-curves for the monaural conditions. The parameter is stimulus type. The effects of a priori probabilities and ears have been averaged.	43
7.	Probability of a given confidence rating, $P(R_j)$, in the binaural conditions with direction, as a function of scale values, R_j. The parameter is stimulus type. The effects of ears and a priori probabilities have been averaged.	47
8.	Type-II ROC-curves for the binaural conditions with direction. The parameter is stimulus type. The effects of a priori probabilities and ears have been averaged.	48
9.	Formant frequency regions of the steady-state vowel stimulus /a/.	65
10.	Formant frequency regions of the vowel-shaped noise stimulus VN.	66

Figure Page

11. Formant frequency region of the narrow-band noise stimulus NBN. 67

12. Schematic diagram of equipment used in generating the six test tapes. 69

13. Distribution of onset time differences between channel A and channel B, in msec, measured with reference to channel A. 73

CHAPTER I

INTRODUCTION

In dealing with binaural interaction in the hearing mechanism in response to complex stimuli, it might be concluded that this mechanism responds to stimuli as meaningfully as it can. For example, the hearing mechanism matches similar sounds for time cues or separates one voice from others speaking at the same time in what has commonly been called the "cocktail-party effect" (Cherry, 1953; Cherry and Taylor, 1954; Kaiser and David, 1960). In addition, it mostly ignores multiple reflections of sound in a room in what has been called the "precedence effect" (Wallach, Newman, and Rosenzweig, 1949). Yet, when presented with quite different sounds at the two ears, it yields two or more images in different locations or sequentially in time, thereby exhibiting a failure to impose order upon the input. It is this latter point which is of primary interest in this paper.

Broadbent (1954) reported an ear effect when two different sets of spoken digits were presented simultaneously to the two ears at the rate of one pair per half second. The listeners tended to report all the digits in one ear before reporting any from the other ear. However, at one pair per second, the listeners tended to report digits temporally (i.e., listeners alternated ears, reporting stimuli delivered to one ear and then the other). In developing a model for this type of behavior, Broadbent (1957, 1958) hypothesized that the ear order effect resulted from listeners adopting the best strategy for retrieving information

presented at rapid rates. The strategy of reporting from one ear first was best because it was assumed that switching attention between the two ears was difficult. In Broadbent's model, the listeners were assumed to attend to one ear while a "filter mechanism" shunted information presented simultaneously to the other ear into short-term storage, for recall later.

In free recall procedures using different stimuli, such as spoken digits or words presented simultaneously to the two ears, several authors have reported an overall right ear superiority.[1] This procedure of simultaneously presenting different stimuli to the two ears has been referred to as dichotic presentation. The procedure generally followed is that of simultaneously presenting different strings of stimuli and requiring the listeners to recall as many of the stimuli as they can. The number of stimuli per string and the rate of presentation vary between experiments. In general, the right ear has been shown to be superior in recall to the left ear in right-handed listeners, in the sense that more stimuli to the right ear are recalled. Kimura (1963b) uses these data to support an "ear asymmetry hypothesis," even though it is probably not the ear to which the asymmetry applies, but rather to some subsystem of the brain. Kimura attributes the "right ear superiority" for speech signals to a greater representation of crossed auditory connections to the left hemisphere.

Since the time of Broca, it has been speculated that one hemisphere of man's brain, usually the left, is specialized for language

---

[1]Bryden, 1962, 1963, 1966; Kimura, 1961a, 1961b, 1963a, 1963b; Bartz, et al., 1967a, 1967b; Satz, et al., 1965; and Dirks, 1964.

and speech functions.[2] Semmes and his colleagues (1960) have suggested a complementary role of the other hemisphere for certain non-language functions. This hypothesis has been demonstrated on patients with confirmed unilateral brain lesions and patients undergoing unilateral removal of brain tissue (Milner, 1958, 1962; Kimura, 1964; and Shankweiler, 1966). Consistent disturbances in recall of verbal material were found in patients with lesions of the left temporal lobe, while patients with similar lesions in the right hemisphere failed to show these verbal defects, but were impaired in recognition of non-verbal patterns. The studies, by Shankweiler, Kimura, Milner, and the others mentioned above, seem to indicate that speech and non-speech signals are processed in different parts of the brain. It must be pointed out, however, that strong contralateral projection does not necessarily mean hemispheric dominance. The demonstration of contralateral projection is not sufficient, in and of itself, to account for Kimura's ear asymmetry hypothesis.

The terms "ear advantage," "ear effect," "ear superiority," and "dominant hemisphere" seem to be used synonymously and suggest an anatomical or physiological advantage of one ear over the other--a view which tends to ignore the highly complex neural connections in the auditory nervous system that are symmetrically and bilaterally projected below the cortex with numerous feedback circuits (Fex, 1967; Rasmussen, 1967). While no one can argue that signal information cannot be carried monaurally, neither can the symmetrical and bilateral projections be ruled entirely superfluous.

---

[2]Eisenson, 1947; Nielson, 1946; Wepman, 1951; Milner, 1954; Penfield and Rasmussen, 1950; and Penfield and Roberts, 1959.

In the studies by Broadbent, Bryden, Kimura, etc., the stimuli were always complete words. A further refinement by Shankweiler (1967) and Shankweiler and Studdert-Kennedy (1967) demonstrated an "ear effect" using syllables as stimuli. They reported the results of several studies using nonense syllables in a dichotic situation in which different speech stimuli were presented simultaneously to the two ears. In this case, the listeners responded after each stimulus presentation. The experimental results showed a substantial and significant right ear advantage in terms of correct recall, indicating that even nonsense syllables presented dichotically were processed unequally by the two cerebral hemispheres. Phonemic structure alone was sufficient to show left hemispheric dominance.

The same authors (Shankweiler and Studdert-Kennedy, 1967) found that the extent of the ear advantage seemed to vary with the composition of the competing consonants. The ear advantage was greater for consonant pairs contrasting in two features (i.e., voicing and place of production) than for consonants contrasting in only one feature. They speculate that the ear is sensitive to information contained in the acoustic signal (i.e., transitions) which corresponds to distinct articulatory dimensions. The authors could find no ear advantage using steady-state vowels, either synthesized or natural. They suggest, therefore, that vowels are processed differently than consonants. The failure to find an ear advantage demonstrated that vowels are processed equally well by both hemispheres. Thus, Shankweiler and Studdert-Kennedy concluded that the ear advantage for consonants pertains to their retention (i.e., storage and recall) and also to their recognition (i.e., discriminability).

The body of data from psychoacoustic experiments, concerned with release of masking for speech or with masking level differences (MLD),

has also used simultaneous presentation of speech stimuli to both ears.[3] Masking level differences (MLD's) are calculated in decibels and represent the improvement in detection or intelligibility that results from using two ears instead of one. Dichotic presentation of stimuli in psychoacoustic experiments refers to stimulus paradigms in which the two ears receive different acoustic inputs. For example, a signal may be presented to one ear; the same signal, plus noise, presented to the other ear; or two independent noises are delivered to the two ears. There may also be differences in phase, time of arrival, or intensity at the two ears. Psychoacoustic experiments, especially in MLD studies, have generally regarded both non-speech and speech signals as simple acoustic events and have presented the same speech signal to both ears.

In psychoacoustic experiments, Carhart (1967) has speculated that in the binaural interaction of speech, the process of auditory localization and the understanding of speech are two functions that are "largely independent" of one another. He further states that stimulation to both ears is required to structure auditory space fully (i.e., to localize the source of the stimulus in space), while intelligibility is achievable through one ear alone. According to this view, binaural factors seem to be unessential to the understanding of speech, although they are necessary for three-dimensionalized sound perception.

Carhart states that the human auditory mechanism consists of two neurological channels, one for each ear, which are relatively independent from one another and which are not required to function together. Either of the two self-sufficient channels is capable of carrying the entire

---

[3]Kock, 1950; Cherry and Taylor, 1954; Schubert, 1956; Schubert and Schultz, 1962; Feldman, 1965; Carhart, Tillman, and Johnson, 1966, 1967; Levitt and Rabiner, 1967a, 1967b.

load of intelligibility for speech when the second channel is receiving a signal that lacks relevance or embodies a second message. This point is clearly brought out in an experiment on selective listening with competing sentences (Cherry and Taylor, 1954), and may also be the basis of the so-called "cocktail-party effect" (Cherry, 1953).

Another point, which can be made from psychoacoustic studies, is that no ear effect or ear dominance has been noted. In fact, Weston, Miller, and Hirsh (1965), in an experiment designed to examine "ear dominance," concluded that sizable differences between right and left ears, with respect to the amount of release from masking for speech, were not typical. The fact that Weston, et al. failed to show an ear effect would agree with Carhart's hypothesis (1967) that either ear is sufficient to carry the entire load of intelligibility when the second ear is receiving a signal that lacks relevance, since the speech stimuli were presented monaurally with different interaural combinations of noise.

A crucial point in evaluating the results of experiments using dichotic presentation of speech stimuli revolves around terminology. In the psychological and speech literature, the dichotic paradigm generally involves the use of different speech stimuli to the two ears. In the psychoacoustic literature, on the other hand, stimuli presented to both ears that are the same in all respects (i.e., intensity, frequency, phase, time of arrival, content, etc.), are generally referred to as DIOTIC. Any binaural condition in which the stimuli to both ears are not identical in all respects is called DICHOTIC. The psychoacoustic experiments, however, have traditionally used the same speech stimulus to both ears, varying the phase, signal-to-noise ratio, time of arrival, and intensity. It is clear that the dichotic paradigm, using different speech stimuli to the two ears (i.e., as used in the psychological and

speech literature), involves subtle changes in phase, intensity, and time of arrival. In addition to these cues, and perhaps of more importance to the higher nervous system, this particular paradigm offers contextual and linguistic cues not available in the traditional psychoacoustic experiments.

In an attempt to tie together the ideas of contralateral projection and hemispheric dominance and to explain the ear effect noted in dichotic presentation of different speech materials presented simultaneously to the two ears, Shankweiler (1967) has suggested that presentation of different speech materials to the two ears is apparently a necessary condition for the occurrence of lateral differences in perception of speech in the presence of competing noise. Indeed, this is the case. If the stimuli are presented for recognition monaurally, neither ear has an advantage, however unfavorable the signal-to-noise ratio.

Liberman and his colleagues (1967) have indicated that the analysis of speech follows a different set of rules than those governing the perception of sensory continua. They hypothesize that the sounds of speech represent a special code to the phonemic structure of language. That is to say, the perception of speech sounds does not follow in a one-to-one relationship to the acoustic units of the original message, as would be required if the perception of speech sounds were in the form of a cipher or alphabet. To do so would require the auditory system to follow discrete acoustic events in rapid running speech at a rate of thirty phonemes per second (Orr, Friedman, and Williams, 1965). Psychoacoustic experiments (Miller and Taylor, 1948) have shown that the temporal resolving power of the auditory system is not that great. Thus, Liberman was led

to hypothesize "parallel transmission" (i.e., simultaneous information about two phonemes present at any instant of time in the acoustic signal), since there is a marked lack of correspondence between sound and the perceived phoneme. The speech signal typically does not contain segments corresponding to the discrete phonemes because the formant transitions are, at every instant in time, providing information about two phonemes: the consonant and the vowel (Liberman, Delattre, Cooper, and Gerstman, 1954; Liberman, 1957; and Delattre, Liberman, and Cooper, 1955).

Liberman and his associates further describe two classes of phonemes: those that show a restructuring of their acoustic cues as a function of context and those that show very little restructuring. Liberman calls the restructured phonemes "encoded" and the unrestructured phonemes "unencoded" and suggests that the perception of these two types of phonemes is different. Most of the consonants fall into the restructured or encoded class. The restructuring is primarily concerned with the first and second formant transitions.

Formant transitions refer to the rapid change in frequency and/or amplitude occurring between phonemes. The first formant transition is generally considered a cue for the perception of manner and voicing, while the second formant transition carries information about the place of production. The encoded phonemes tend to be categorical, that is, listeners can discriminate only slightly better than they can absolutely identify. One tends, therefore, to hear quantal jumps from one perceptual category to another. Unencoded phonemes, however, tend to be perceived in a continuous manner. Listeners can, therefore, discriminate many more stimuli than they can absolutely identify (Studdert-Kennedy,

Liberman, and Stevens, 1963, 1964; Stevens, Öhman, Studdert-Kennedy, and Liberman, 1964), since they can follow continuous changes in the signal. Liberman, et al. (1967) suggests that the perception of encoded phonemes differs from that of unencoded phonemes or non-speech and requires the participation of different parts of the brain. Recovery of the encoded phonemes in the speech signal must, he says, require a special decoder; the operation of such a decoder suggests in turn a "speech mode."

Although Shankweiler and Studdert-Kennedy have suggested that the ear advantage pertains not only to retention, but to recognition as well, this concept has not been tested experimentally. All studies using dichotic presentations have required retention and recall on the part of the subject by virtue of the fact that the response always consisted of repeating one or several stimuli. Regardless of whether the listeners heard one stimulus before responding or whether they heard a list of stimuli, the mode of responding (i.e., repeating the stimulus or stimuli) required that the signal be discriminated and retained. It, therefore, also required some degree of memory storage, since recognition must precede retention.

If the speech mode exists for decoding restructured phonemes, as suggested by Liberman, listeners in all previous research in dichotic speech analysis have had the advantage of operating continuously in this mode. Regardless of the procedure followed, listeners expected a speech stimulus on each trial, and, as a result, they did not have to make a decision concerning speech or non-speech prior to neurological processing of the stimuli.

It can be hypothesized that the ear effect will only be found

when: (1) competing speech stimuli are presented simultaneously, and (2) the behavioral response requires stimulus discrimination (i.e., recognition).

The purpose of this study is to test this hypothesis by attempting to answer the following questions: (1) Can the auditory system differentiate between speech and non-speech stimuli at a specified detection level? (2) If the human auditory system can, in fact, differentiate between speech and non-speech stimuli at the level of detection (i.e., audibility threshold), is there an ear effect? (3) Is the ear effect enhanced by simultaneous dichotic presentation?

CHAPTER II

PROCEDURE

Subjects

Ten adult female listeners, without prior training in speech or hearing science and unaware of the purpose of this research, were used as subjects. The listeners ranged in age from 20 to 27 years, with a median age of 21.5 years. All listeners had normal hearing and no more than 5 dB difference in acuity between the ears at any octave frequency between 125 and 8000 Hz, as determined by audiometric testing.

Stimuli

Four different stimuli were presented to the subjects: a synthesized consonant vowel combination (CV), a synthesized steady-state vowel (V), a vowel-shaped noise (VN), and a narrow-band noise (NBN).

Consonant-vowel combination. - The CV stimuli were synthesized and recorded on magnetic tape at the Haskins Laboratory and consisted of an 80 msec consonant, followed by a 220 msec steady-state vowel /a/ with a fundamental frequency of 138 Hz. The first formant transition of the consonant began at 280 Hz and reached a final value of 743 Hz in 80 msec. The first formant transition of the consonant and the steady-state portion of the vowel were the same for all stimuli. The second and third formant frequency regions of the consonant transition were variable. F_2 consisted of eleven initial values ranging from 543 Hz to 2837 Hz, and F_3 consisted of eight initial values ranging from 1190 Hz to 3530 Hz. The

formant frequencies of the consonant were combined in such a way that the CV stimuli resembled voiceless fricatives in the general area between /f/ and /θ/. A complete description of the CV stimulus and the other stimuli can be found in Appendix A.

Steady-state vowel. - The steady-state vowel stimulus /a/ was synthesized on a passive vowel synthesizer at the Communication Sciences Laboratory of the University of Florida, using a saw-tooth wave with a fundamental frequency of 138 Hz as the voice source. The vowel stimulus contained the same formant frequency energy concentrations and fundamental frequency as the vowel in the CV combination, with the formant regions and relative amplitudes essentially those reported by Peterson and Barney (1952). The output of the synthesizer was recorded continuously on magnetic tape.

Vowel-shaped noise. - With the passive vowel synthesizer adjusted for the formant regions of the vowel /a/, white noise was used to drive the synthesizer rather than the saw-tooth wave. This resulted in a noise stimulus containing energy concentrations resembling the formant structure of the vowel /a/. The output of the synthesizer was also recorded continuously on magnetic tape.

Narrow-band noise. - The narrow-band noise stimulus consisted of the F_1 region of the vowel-shaped noise. The F_1 region was chosen simply because it was representative of at least one aspect of the vowel. The narrow-band noise was also recorded continuously on magnetic tape.

Thus, the stimuli, used in this study, consisted of two samples of synthesized speech (CV and V) and two samples of non-speech (VN and NBN). The CV stimulus resembles the "encoded phoneme" described

by Liberman, et al. (1967), while the V stimulus resembles the "unencoded phoneme." The two non-speech stimuli, while consisting of certain spectral characteristics common to the speech stimuli, are considered simply as noise stimuli.

Test Stimuli

Six two-track tape recordings were generated from the three original stimulus recordings (see Appendix A) on an Ampex PR-10/2 tape recorder. Each recording contained 115 stimuli on each track. The position and number of occurrences of each of the four stimulus types (CV, V, VN, NBN) on each track of the test tapes were determined randomly. The onset time of competing stimuli on the two tracks was identical, except as noted in Appendix B. In addition, the rise-fall time (10 msec) and the duration (300 msec) of all stimuli were identical. The test tape recordings were generated in such a way that on two tapes the a priori probability of occurrence of speech stimuli was 0.75. On two other tapes, the probability of occurrence of speech was 0.50; and, for another two tapes, speech occurred with a probability of 0.25. Test tapes with different a priori probabilities for the occurrence of speech stimuli were used to prevent the subjects from expecting a speech stimulus on each presentation. A detailed description of the generation of the test tapes can be found in Appendix B.

The responses to the first fifteen stimuli on each test tape were used in obtaining the detection level and served as practice stimuli. These stimuli, in addition to an initial one-hour practice session, familiarized the listeners with the task.

Experimental Apparatus

Figure 1 shows a schematic diagram of the equipment used in each experimental session. The test tapes were played from the two-channel tape recorder (Ampex PR-10/2) on which the stimuli were originally recorded. Two 10 dB attenuators, one for each channel, were used to compensate for small differences in output voltage level between the channels. An event counter, located inside the sound-treated room, was triggered from one channel of the test tape through an oscilloscope (Tektronix, Model 531). The counter served two purposes. First, it indicated the number of the next stimulus, a procedure which was deemed necessary since the listeners were working at or near the detection threshold and, therefore, would be expected to miss some of the stimuli. The indication on the counter enabled the listeners to record their response to the appropriate stimulus. Second, the counter served as a warning signal to the listeners since it indicated the number of the next stimulus three seconds prior to its presentation.

The output voltages of channels A and B of the tape recorder were adjusted by two attenuators (Hewlett-Packard, Model 350D), and the signals were fed to two earphones (Telephonics, TDH-39) through impedance matching transformers (United Transformer Corporation, LS-33).

Experimental Session

The listeners were tested individually, while comfortably seated within a double-walled, sound-treated room (Industrial Acoustics Corporation, Model 1204). The counter was placed immediately in front of the subject so it could be viewed at all times, and its operation explained.

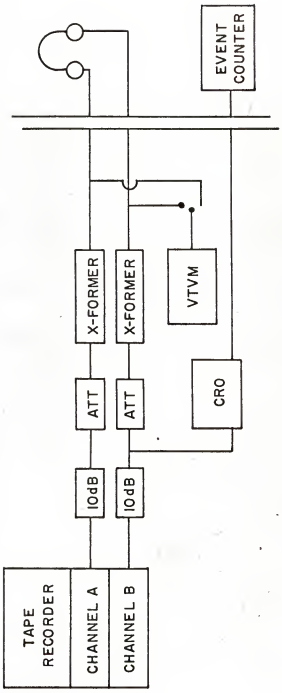

Figure 1. Schematic diagram of equipment used in each experimental session

The 50 percent detectability threshold (monaural) for the initial fifteen stimuli of all test tapes was derived using the method of limits from each ear of each listener (with a separate recording channel for each ear). This level of presentation was selected for all stimuli since one objective of the experiment was to have the listeners operate as closely as possible at a specified detection level. The detectability threshold was derived monaurally, as unequal sensitivity between the two ears could bias the results in the direction of a more sensitive ear and, thus, lead to an artifactual ear effect.

The listeners were given a form on which to record their responses. The form was numbered from 1 to 115, with four headings beside each number. The headings represented speech, non-speech, no stimuli (if not heard), and a confidence rating of each speech or non-speech judgment (discussed below).

The listeners were instructed simply to listen for the presentation of a stimulus and to make a decision based on their best guess as to whether the stimulus was speech or non-speech.

All listeners participated in all experimental conditions. For each of the three a priori probabilities for the occurrence of speech, the listeners responded monaurally in the right and left ears; binaurally, with instructions to record their responses on the basis of stimuli presented only to the right ear; binaurally, with instructions to respond on the basis of stimuli presented only to the left ear; and binaurally, with no directions as to ear. All conditions were randomized.

Confidence Ratings

It stands to reason that a listener has more information available than the single response recorded in an identification test, which may

only be scored as correct or incorrect. The listener must also have a certain amount of confidence in this binary response. Sometimes, the response may simply be a guess; and other times, the listener may be quite sure. Clarke (1960) has demonstrated that confidence ratings carry information independent of the listener's original identification. It has also been shown by Pollack and Decker (1958) that the addition of the confidence rating does not interfere with the listener's identification score.

Confidence judgments have been studied by means of Receiver Operating Characteristics (ROC) curves (Egan, 1957). Type-II ROC-curves are commonly used. This type of curve shows the probability that a receiver will confirm the response, given the response is correct, against the probability that the receiver will confirm the response when, in fact, it is incorrect.

The listeners, of course, have four possible responses they can make. First, they can say they are correct when, in fact, they are correct (P(Y/C)). Second, they can say they are correct when, in fact, they are incorrect (P(Y/I)). Third, they can say they are incorrect when, in fact, they are correct (P(N/C)). And, fourth, the listeners can say they are incorrect when, in fact, they are incorrect (P(N/I)). A plot of the first two responses (i.e., P(Y/C) and P(Y/I)), the familiar "hits" and "false alarms" of Signal Detection Theory, generates the ROC-curves.

Another measure available from confidence judgments is the accuracy of attaining optimum values. The optimum criterion value can be read directly from normal-normal probability paper by finding the point where the ROC-curve crosses a negative diagonal drawn from the corner of the

graph to the 50 percent point of the chance line. The Type-II ROC-curves were derived originally by using a binary-decision procedure. However, Pollack and Decker (1958) and Clarke (1960) have shown that the Type-II ROC-curve can be derived from ratings.

The listeners in this study were asked to make confidence judgments of their responses on a five-point scale. The scale used in the confidence judgments follows:

5 - Positive best guess is correct

4 - Fairly sure best guess is correct

3 - Cannot decide

2 - Have little confidence in best guess

1 - Have no confidence in best guess

The listeners made a confidence rating after each stimulus presentation.

CHAPTER III

RESULTS

The ten listeners participated in all fifteen experimental conditions. For each of the three a priori probabilities for the occurrence of speech (i.e., 0.25, 0.50, and 0.75), stimuli were presented to the right and left ears monaurally; binaurally, with directions to respond only on the basis of the stimuli in the right ear; binaurally, with directions to respond only on the basis of the stimuli in the left ear; and binaurally, with no specific directions. The listeners' task was to make a decision, based on their best guess, as to whether a particular stimulus was speech or non-speech and to rate their response, based on the amount of confidence they had in the correctness of the response.

Since the data were recorded in the form of proportions, they were submitted to an arcsin transformation prior to statistical analysis (Walker and Lev, 1953). This technique is useful because the arcsin distribution is approximately normal with unit variance. Homogeneity of variance cannot be assumed with proportions because the variance is determined by both the proportion and the total number of presentations (N). If all proportions are based on the same N and if each is transformed to an arcsin, homogeneity of variance is then assured because each arcsin has the same variance, 1/N, even though the proportions differ. The summary tables for all statistical treatments can be seen in Appendix C.

Monaural Condition

Ear Effect

The percentage of correct responses for the four different stimuli in the monaural conditions can be seen in Figure 2 and Table 1. As stated above, an arcsin transformation was required before statistical analyses. Results of t-tests (Edwards, 1950) and analyses of variance (Winer, 1962) on the transformed data failed to show a significant difference between the two ears for any of the four stimuli or any of the three a priori probabilities. Summary tables are found in Appendix C. Figure 2, then, is based on the mean percentage correct scores taken over the three a priori probabilities shown in Table 1. It can be seen that the CV stimulus was generally detected as speech, and the two noise stimuli were generally detected as non-speech. The V stimulus, however, was detected both as speech and non-speech in both ears.

Effects of Stimulus Types

An analysis of variance was carried out on the transformed percentage correct scores in order to determine the effects of the different stimuli. The results showed a significant difference due to stimulus type. Scheffé's a posteriori test for the comparison between means (Winer, 1962) was carried out to determine the source of this variation. Results of this test showed the V stimulus to be statistically different from the other three stimuli. These results might be expected from Figure 2, where the V stimulus was considerably lower in percent correct than the other three stimuli. The summary tables for the analysis of variance and the results of the a posteriori comparisons between the means can be seen in Appendix C.

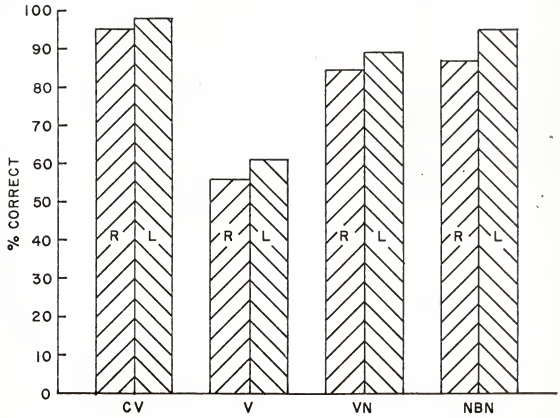

Figure 2. Mean percent correct responses for right and left ears in the monaural listening conditions for the four stimulus types. The values are averaged over a priori probabilities

Table 1. Percent correct responses for right and left ears in the monaural condition for each of the three a priori probabilities

P(Speech)	Stimuli	Right Ear	Left Ear
0.25	CV	95.85	98.90
	V	57.50	62.75
	VN	76.25	93.75
	NBN	83.00	97.60
0.50	CV	97.30	97.45
	V	66.50	63.50
	VN	88.75	94.75
	NBN	93.50	94.45
0.75	CV	93.90	96.05
	V	46.00	59.00
	VN	89.25	77.00
	NBN	86.25	90.90
Mean	CV	95.75	97.60
	V	56.50	61.50
	VN	85.00	89.50
	NBN	87.50	95.75

Effects of Probability for the Occurrence of Speech

An analysis of variance was carried out to test the effect of a priori probability for the occurrence of speech. Results of this analysis failed to demonstrate a statistically significant effect. The summary table for this analysis can be seen in Appendix C.

Stimulus Omissions

Since the listeners were operating close to the threshold of audibility, it was expected that they would omit (i.e., not hear) some of the stimuli. While this was the case, there were no apparent differences in the number of omissions from session to session. The listeners, on the average, omitted 29.5 percent of the stimuli during each experimental session in the monaural conditions (Table 2). This indicates that the listeners were highly motivated and consistent in their responses and that they were operating fairly close to the threshold of detectability. It is of interest to note that the CV stimulus was omitted, on the average, fewer times than any of the other three stimuli. This behavior is not easily explained since the differences in peak intensity between all stimuli was only ± 1.4 dB (see Appendix B).

Response Errors

The average number of errors made during an experimental session for the three different a priori probabilities of speech occurrence can be seen in Table 3. The listeners made, on the average, 14.2 errors in the right ear and 8.6 errors in the left ear. The largest number of errors were made in those conditions where the a priori probability for the occurrence of speech was 0.75. This might be expected since the largest number of errors were made for the V stimulus, and this stimulus occurred more often at that probability.

Table 2. Average number of omissions for each stimulus type during an experimental session of 100 stimuli. The omissions are averaged over the three a priori probabilities

Experimental Condition	Stimulus Type				
	(CV)	(V)	(VN)	(NBN)	Total
Monaural					
Right	1.9	9.1	11.0	6.0	28.0
Left	1.9	10.4	12.3	7.3	31.9
Binaural (with direction)					
Right	3.5	10.8	12.8	8.8	35.9
Left	4.0	12.2	11.8	8.0	36.0

Table 3. Average number of errors during an experimental session for the three a priori probabilities of speech occurrence and the different experimental conditions

Experimental Condition	A Priori Probability			
	0.25	0.50	0.75	Mean
Monaural				
Right	13.4	13.1	16.1	14.2
Left	6.5	8.2	11.1	8.6
Binaural (with direction)				
Right	9.6	13.1	15.6	12.8
Left	15.1	16.2	16.3	15.9

Summary of Results in Monaural Conditions

In the monaural conditions, no significant differences in performance were found between the right and left ears. In addition, there were no significant effects due to a priori probabilities for the occurrence of speech. There was, however, a significant difference between the stimulus types. The CV stimulus was generally reported as speech, while the two noise stimuli (VN and NBN) were generally detected as non-speech. The V stimulus, on the other hand, was detected equally as speech and non-speech. The CV stimulus was omitted fewer times than the V stimulus, VN stimulus, or NBN stimulus, although the peak voltage of the four stimulus types were approximately equal.

Binaural Conditions with Directions

Ear Effect

In this condition, the listeners were instructed to respond only on the basis of the stimulus presented in the right or left ear, although a competing stimulus was presented simultaneously in the opposite ear. The percentage of correct responses for the four different stimuli can be seen in Figure 3 and Table 4. Results of t-tests and analyses of variance on the transformed data failed to demonstrate a significant difference between the two ears or any of the three a priori probabilities. Summary tables are found in Appendix C. These results fail to agree with previous research which generally shows an ear effect when speech stimuli are presented dichotically.

Effects of Stimulus Types

An analysis of variance was carried out on the transformed percentage correct scores in order to determine the effects of stimulus

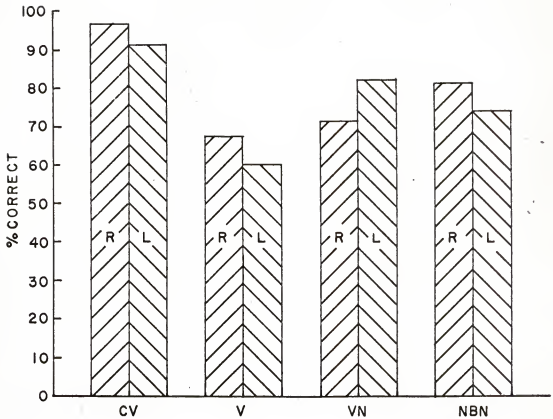

Figure 3. Mean percent correct responses in right and left ears for the four stimulus types in the binaural conditions with direction. The values are averaged over a priori probabilities

Table 4. Percent correct responses for right and left ears in binaural condition with direction for each of three *a priori* probabilities

P(Speech)	Stimuli	Right Ear	Left Ear
0.25	CV	99.35	87.00
	V	74.00	52.50
	VN	77.50	83.00
	NBN	90.00	81.90
0.50	CV	95.85	85.75
	V	65.75	64.90
	VN	74.50	77.50
	NBN	76.25	77.75
0.75	CV	92.25	95.60
	V	61.90	62.90
	VN	60.50	88.50
	NBN	76.25	63.50
Mean	CV	96.45	90.50
	V	67.50	60.00
	VN	71.00	82.50
	NBN	81.50	74.50

type. The results of this analysis showed a significant difference due to stimulus type, and Scheffé's a posteriori test, for comparison between means, showed the CV stimulus significantly different from the other three. The summary table for the analysis of variance and the results of the a posteriori comparisons can be seen in Appendix C. The results of the analysis of variance might be expected from the data in Figure 3, where the percent correct scores for the CV stimulus are considerably higher than for the other three stimuli. It is interesting to note that the CV stimulus produces the significant difference between stimulus types in the binaural conditions with directions, while the V stimulus seems to produce this difference in the monaural conditions. In viewing Figure 2 and Figure 3, it can be seen that the CV stimulus has not changed in percent correct, but rather, the two noise stimuli have decreased in the binaural conditions with directions. This point will be discussed further in the section on "Response Errors."

Effects of Probability for the Occurrence of Speech

An analysis of variance was carried out to determine the effects of the different a priori probabilities in the binaural conditions where the listeners were instructed to report from a particular ear. The results of this analysis showed that a priori probability for the occurrence of speech did not produce a significant effect. The summary table for this analysis can be seen in Appendix C.

Stimulus Omissions

The listeners averaged omitting 35.95 stimuli in each experimental session in the binaural conditions with directions (see Table 2), with no apparent differences in number of omissions from session to session.

It is evident from Table 2 that the CV stimulus was omitted, on the average, fewer times than the other three stimuli, as was noted in the monaural conditions. Again, this behavior is not easily explained in light of the approximately equal peak intensity values of the different stimulus types.

Response Errors

The percentage of response errors in the reporting ear, when competing stimuli is presented simultaneously to the opposite ear, are reported in Table 5. It can be seen from this table (right column) that 14.0 percent of the errors were due to calling a CV stimulus non-speech; 35.7 percent of the errors were due to calling a V stimulus non-speech; 22.3 percent of the errors were due to calling a VN stimulus speech; and 28 percent of the errors were due to calling a NBN stimulus speech. On the other hand, 35.3 percent of the errors in the reporting ear were made when a CV stimulus was present in the contralateral ear (bottom row); 23.1 percent of the errors were made when a V stimulus was present in the contralateral ear; 21.7 percent of the errors when a VN was presented in the contralateral ear; and 19.9 percent of the errors when a NBN stimulus was present in the contralateral ear.

The largest number of errors occurred when a VN stimulus or a NBN stimulus in the reporting ear was in competition with a CV stimulus in the contralateral ear (11.7 percent and 16.8 percent, respectively). In other words, the listeners seemed to be attending to the CV stimulus, regardless of instruction to report from a particular ear. This apparent tracking of the CV stimulus resulted in a greater number of errors being made for the VN and NBN stimuli and, thus, the decrease in percent correct seen in Figure 3. It can also be seen from Table 5 that very

Table 5. Response errors in the reporting ear (in percent) when competing stimuli are presented in the contralateral ear

Reporting Ear	Contralateral Ear				Total Errors in Reporting Ear
	CV	V	VN	NBN	
CV	2.2	2.0	4.6	5.2	14.0
V	4.6	10.9	11.7	8.5	35.7
VN	11.7	5.3	3.2	2.1	22.3
NBN	16.8	4.9	2.2	4.1	28.0
Total Errors	35.3	23.1	21.7	19.9	

few errors were made when a noise stimulus in the reporting ear was presented in competition with another noise stimulus in the contralateral ear. In other words, when noise stimuli were presented to both ears simultaneously, the listeners had little trouble detecting the stimulus as non-speech. By the same token, very few errors were made when a CV stimulus in the reporting ear was in competition with another CV stimulus in the contralateral ear (2.2 percent). This is not true, however, for the V Stimulus, which shows the greatest percentage of errors in the reporting ear (35.7 percent), and a fairly even distribution of errors in competition with the different stimulus types presented in the contralateral ear.

Summary of Results in Binaural Conditions with Directions

In the binaural conditions with directions, no significant differences in performance were found between the right and left ears. There were no significant effects due to a priori probability for the occurrence of speech. There was, however, a significant difference between stimulus types. As in the monaural conditions, the CV stimulus was generally reported as speech, while the two noise stimuli (VN and NBN) were generally detected as non-speech. There was, however, an increase in the errors for the two noise stimuli (i.e., calling these stimuli speech) when they were in competition with a CV stimulus in the contralateral ear.

Binaural Conditions without Directions

Ear Effect

In this condition, the listeners were instructed to listen for a stimulus and make a decision as to whether it was speech or non-speech.

They were given no instructions pertaining to the ear in which they should attend. The percentage of correct responses for the four different stimuli can be seen in Figure 4 and Table 6. Results of t-tests on the transformed data failed to demonstrate a significant difference between the two ears. These tables may be seen in Appendix C. A striking feature of Figure 4 is the decrease in percentage of correct responses for the VN stimulus and the NBN stimulus. The percentage of correct responses for the CV stimulus, however, is consistent with those obtained for the monaural conditions and the binaural conditions with directions.

Effects of Stimulus Types

An analysis of variance was carried out to assess the effects of different stimulus types in this experimental condition. The summary table for this analysis can be seen in Appendix C. The results of this analysis indicated a significant difference due to stimulus type. The Scheffé test for a posteriori comparisons between means was performed, and indicated the difference was due to the CV stimulus. The summary table for this a posteriori comparison can be seen in Appendix C. Again, these results might be expected from the data in Figure 4, which show the percentage of correct responses for the CV stimulus to be considerably higher than the other three stimulus types.

Effects of Probability for the Occurrence of Speech

An analysis of variance was carried out to determine the effect of the three different a priori probabilities of speech occurrence. The summary table for this analysis can be seen in Appendix C. The results showed the effects of different a priori probabilities significant at the 0.05 level of confidence. The Scheffé test for a posteriori

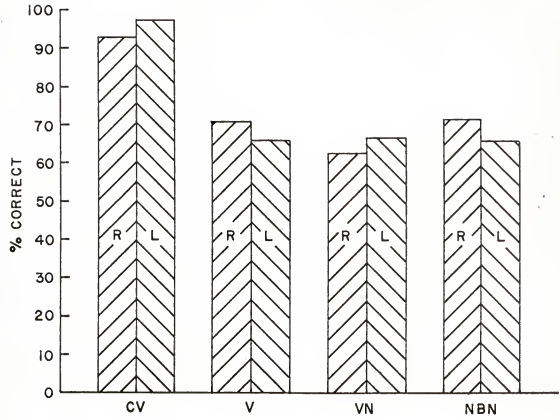

Figure 4. Mean percent correct responses in right and left ears for the four stimulus types in the binaural conditions without direction. The values are averaged over a priori probabilities.

Table 6. Percent correct responses for right and left ears in the binaural condition without direction for each of three *a priori* probabilities

P(Speech)	Stimuli	Right Ear	Left Ear
0.25	CV	94.90	99.10
	V	65.75	44.00
	VN	79.90	81.75
	NBN	82.75	84.75
0.50	CV	93.25	93.90
	V	69.00	62.50
	VN	57.50	50.50
	NBN	75.50	71.50
0.75	CV	93.00	98.05
	V	77.90	72.25
	VN	49.50	65.90
	NBN	51.50	37.25
Mean	CV	93.50	97.45
	V	71.00	66.00
	VN	62.50	66.50
	NBN	70.50	65.50

comparisons between means was carried out, and indicated the source of variation was caused by the large differences between mean values between the a priori probabilities of 0.25 and 0.75. This difference is not easily explained in light of the failure to find differences in a priori probabilities in the other experimental conditions.

Response Errors

An analysis of response errors was performed as though the listeners were reporting from the right ear only and, again, as though they were reporting from the left ear, in relation to the stimulus presented in the contralateral ear. As no ear effect was present, these data were averaged and can be seen in Table 7. It can be seen from the table (right column) that 7.7 percent of the errors in the reporting ear were due to calling the CV stimulus non-speech, 35.3 percent due to calling the V stimulus non-speech, etc. On the other hand, 40.9 percent of the errors in the reporting ear were made when a CV stimulus was present in the contralateral ear (bottom row). As in the binaural conditions with directions, the largest number of errors occurred when a VN stimulus or a NBN stimulus in the reporting ear was in competition with a CV stimulus in the contralateral ear (18.9 percent and 20.0 percent, respectively). It is also obvious that very few errors were made when a noise stimulus in the reporting ear was in competition with another noise stimulus in the contralateral ear. By the same token, very few errors were made when a CV stimulus was in competition with another CV stimulus (0.2 percent).

Summary of Results in Binaural Conditions without Directions

In the binaural conditions without directions, no significant differences were found between the right and left ears. A significant

Table 7. Response errors in the binaural without direction condition (in percent)

Reference Ear	Contralateral Ear				Total Percent in Reference Ear
	CV	V	VN	NBN	
CV	0.2	1.8	2.0	3.7	7.7
V	1.8	10.6	12.0	10.9	35.3
VN	18.9	8.6	1.1	1.0	29.6
NBN	20.0	5.7	1.0	0.7	27.4
Total Percent	40.9	26.7	16.1	16.3	

effect due to a priori probabilities for the occurrence of speech was apparent. There was also a significant effect due to stimulus types. As in the monaural conditions and the binaural conditions with directions, the CV stimulus was generally reported as speech. There was, however, an increase in response errors for the two noise stimuli (i.e., calling noise stimuli speech), over the binaural conditions with directions, when they were in competition with a CV stimulus in the contralateral ear.

Confidence Ratings

In addition to making a decision as to whether the stimulus detected was speech or non-speech, the listeners were instructed to rate the amount of confidence they had in their decision on a five-point scale. The confidence ratings were analyzed for the monaural conditions and the binaural conditions with directions.

Monaural Conditions

ROC-curves, derived from the rating scale values, were plotted for the six monaural conditions (i.e., the three a priori probabilities in the right and left ears), and the optimum values for each of these conditions obtained. The optimum value is the point on the ROC-curve where the ratio of correct responses, or "hits," to errors, or "false alarms," is optimum. The optimum value, then, is the point where the listeners maximize their "hits" and minimize their "false alarms." The optimum values for the monaural conditions can be seen in Table 8. There seems to be no consistent tendency for optimum values to increase with increased a priori probability for the occurrence of speech.

The mean optimum values averaged over different a priori probabilities and the standard deviations of the mean for the different

Table 8. Optimum values computed for the monaural right and monaural left conditions

Stimulus Type	Right Ear			Left Ear		
	0.25	0.50	0.75	0.25	0.50	0.75
CV	.48	.69	.71	.62	.82	.72
V	.39	.48	.47	.74	.40	.63
VN	.40	.41	.61	.62	.69	.46
NBN	.50	.70	.69	.82	.65	.48

Table 9. Mean optimum values and standard deviations for four stimulus types in the monaural right and monaural left conditions. The effects of a *priori* probability of speech have been averaged

Stimulus Type	Right Ear		Left Ear	
	Mean	S D	Mean	S D
CV	.63	.10	.72**	.08
V	.44	.04	.55	.14
VN	.47	.10	.59	.10
NBN	.63	.10	.65	.14

**Significantly greater than chance at the 0.01 confidence level.

stimulus types can be seen in Table 9. In only one stimulus condition, the CV stimulus in the left ear, did the listeners' confidence in their responses differ significantly from chance. In other words, while the listeners, in general, could perform the experimental task, they generally had little confidence in their performance, except when CV stimuli were presented to the left ear.

Figure 5 shows the probability of a given response category (i.e., particular rating, R_j) to occur for each of the four stimulus types under monaural conditions. It can be seen that more ratings of 3, 4, and 5 are given in response to CV and NBN stimuli than to V and VN stimuli. In addition, more ratings of 1 and 2 were given to the VN and V stimuli than to the CV and NBN stimuli. In other words, the listeners seem to be more confident in their responses at the two extremes: the CV stimulus, which most resembles speech, and the NBN stimulus, which is probably the least like speech.

Figure 6 shows the ROC-curves for the monaural conditions. Since the experimental conditions involving a priori probabilities and ears (right vs. left) did not produce significant effects, the data were combined in this curve. Only the curve representing the CV stimuli differs statistically from the chance line, thus illustrating the results based on optimum values discussed above.

Binaural Conditions with Directions

The optimum values obtained for the binaural with direction condition can be seen in Table 10. No consistent pattern is evident. The mean optimum values, averaged over different a priori probabilities, and standard deviations of the mean for the different stimulus types in the two ears, can be seen in Table 11. Here, the listeners' confidence

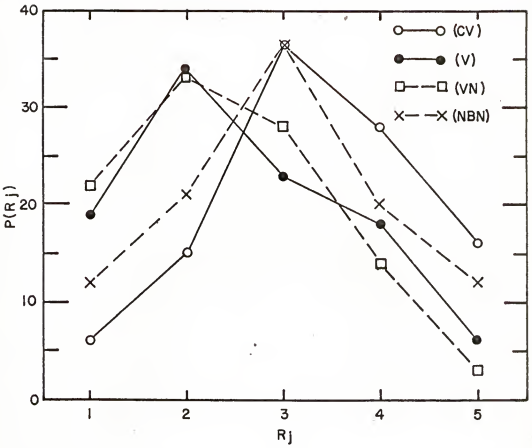

Figure 5. Probability of a given confidence rating, $P(R_j)$, in the monaural conditions as a function of scale values, R_j. The parameter is stimulus type. The effects of ears and a priori probabilities have been averaged.

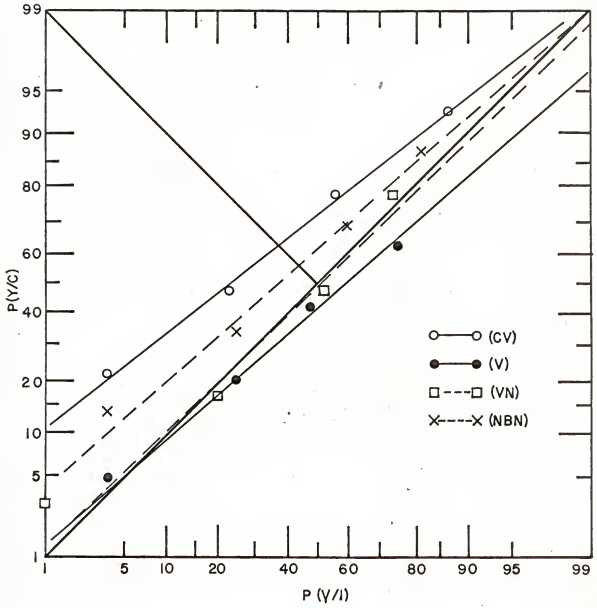

Figure 6. Type-II ROC-curves for the monaural conditions. The parameter is stimulus type. The effects of a priori probabilities and ears have been averaged.

Table 10. Optimum values computed for the binaural with direction condition

Stimulus Type	Right Ear			Left Ear		
	0.25	0.50	0.75	0.25	0.50	0.75
CV	.78	.81	.61	.62	.58	.72
V	.50	.50	.62	.45	.49	.52
VN	.60	.50	.42	.51	.43	.57
NBN	.67	.57	.57	.67	.54	.52

Table 11. Mean optimum values and standard deviations for the four stimulus types in the binaural with direction condition. The effects of a priori probability of speech have been averaged

Stimulus Type	Right Ear		Left Ear	
	Mean	S D	Mean	S D
CV	.73**	.09	.64**	.06
V	.54	.06	.48	.03
VN	.51	.07	.50	.06
NBN	.60*	.05	.58	.07

*Significantly greater than chance at the 0.05 confidence level.

**Significantly greater than chance at the 0.01 confidence level.

that a CV is, in fact, a speech stimulus is better than chance in both ears with respect to the CV stimuli. With all other stimulus conditions, except NBN in the right ear, little confidence in performance is indicated by the listeners.

Figure 7 shows the probability of a given response category, $P(R_j)$, occurring for each of the four stimulus types under the binaural conditions with directions as a function of scale value (R_j). In the binaural condition, a larger number of 1 ratings and fewer 2 ratings were used for the V stimulus than were used in the monaural conditions. Also, the listeners tended to use more 3 ratings for the VN stimulus in the binaural conditions. The CV and NBN stimuli continued to show a greater number of 4 and 5 ratings.

Figure 8 shows the ROC-curves for the binaural conditions with directions, since the experimental conditions involving a priori probabilities and ears (right vs. left) failed to produce significant effects. These data were combined in this curve. In the figure, the curves representing the CV and NBN stimuli differ significantly from the chance line, illustrating the results based on the optimum values in Table 11.

Summary of Results of Confidence Ratings

The confidence ratings demonstrated that while the listeners could make a decision of speech or non-speech in response to stimuli, they were generally not confident in their response, except when presented with the CV stimulus.

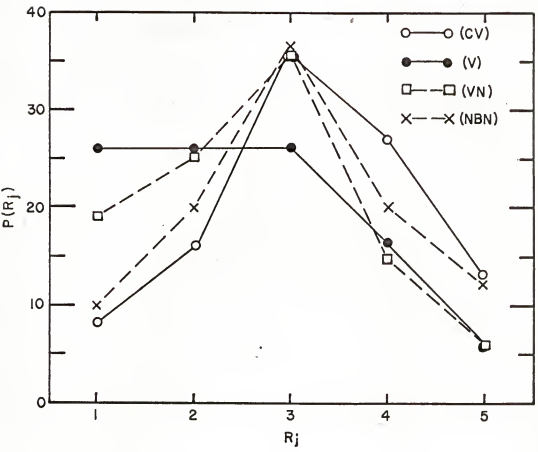

Figure 7. Probability of a given confidence rating, $P(R_j)$, in the binaural conditions with direction, as a function of scale values, R_j. The parameter is stimulus type. The effects of ears and *a priori* probabilities have been averaged.

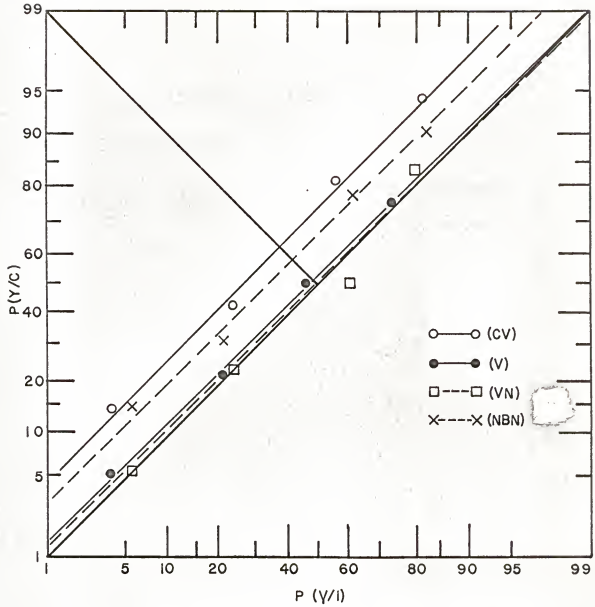

Figure 8. Type-II ROC-curves for the binaural conditions with direction. The parameter is stimulus type. The effects of a priori probabilities and ears have been averaged

CHAPTER IV

DISCUSSION

Speech Versus Non-Speech

One of the purposes of the present research was to determine if listeners could differentiate between speech and non-speech at a specified detection level. Results of this investigation indicate that listeners are able to make such a distinction with a high degree of accuracy. This general statement seems to hold particularly for the CV stimulus and for the two noise stimuli (VN and NBN). The CV stimulus was detected as speech a high percentage of the time (96 percent), and the two noise stimuli were correctly identified as non-speech (87 and 91 percent, respectively). The V stimulus, however, was detected both as speech and non-speech.

In previous research using dichotic presentation of speech material, the listeners expected a speech stimulus of each presentation. It would seem reasonable to assume that the listeners' performance might be influenced by this factor. If a special speech mode exists, as suggested by Liberman, et al. (1967), listeners have had the advantage of continually operating in this mode, since a speech stimulus had been presented on each trial. In order to eliminate this advantage, the present research used different a priori probabilities for the occurrence of speech, which were unknown to the listeners. The results of this study indicated that different a priori probabilities were not a significant factor in

differentiating between speech and non-speech. This would seem to indicate that the auditory system is extremely efficient as a detector, and that a probability for the occurrence of speech of 1.00 is not essential in making the differentiation between speech and non-speech.

Ear Effect

Previous research has failed to demonstrate an "ear effect," that is, general superiority of one ear-brain system over the other, when stimuli were presented monaurally (see Dirks, 1964; Shankweiler, 1967; Shankweiler and Studdert-Kennedy, 1967). In this respect, the present study agrees with previous research. At the 70 percent detection level, the ten listeners in this experiment failed to exhibit a superiority of one ear-brain system over the other.

The aforementioned studies, and others (Kimura, 1961a, 1961b, 1963a, 1963b; Bryden, 1962, 1963, 1966; etc.), however, have demonstrated a consistent and significant right ear advantage in right-handed listeners, when stimuli were presented dichotically. The listeners in previous experiments have demonstrated greater recognition of stimuli presented to the right ear than to the left ear. At the detection level used in the present study, no such ear-brain superiority was confirmed for any of the dichotic experimental conditions.

The common explanation of the ear-brain superiority has been from anatomical and physiological bases. Kimura (1963b) attributed a "right ear superiority" for speech signals to a greater representation of crossed auditory connections to the left hemisphere. However, there is no reason to assume that one ear would be superior to another on the basis of neural projections. Electrophysiological studies in

animals, using gross electrodes (Woolsey and Walzl, 1942; Tunturi, 1946; Rosenzweig, 1951) and microelectrode and single unit responses (Gerstein and Kiang, 1964; and Hall and Goldstein, 1968) do demonstrate that contralateral projections (i.e., right ear to left cortex) are the primary pathways, but they also demonstrate that both ears are represented in a single hemisphere. Moreover, Gacek (1967) demonstrates that the auditory nervous system below the cortex, while predominately crossed, is symmetrical and approximately the same for both ears. There seems to be no greater number of fibers in one ear-brain system than in the other.

If greater representation of neural fibers could be shown for the right ear than the left ear, for example, one could quite easily predict a right ear-brain superiority for speech presented monaurally to the right ear. Such neural representation has not been found to exist. In addition, experimental data resulting from monaural stimulation do not show ear superiority. The present research confirms the absence of ear superiority for monaurally presented stimuli.

The fact remains, however, that ear superiority has been well demonstrated in the past with dichotically presented speech stimuli. The question might be asked, then, why no ear effect from binaural stimulation was found in the present experiment. The answer may be due to the fact that in all previous research on dichotic presentation of speech, the stimuli have been presented at sound pressure levels well above the detection level. The listeners, in the present study, were unaware of the precise nature of the speech stimuli presented to them and, for that matter, if <u>any</u> speech signal was presented. In the previous research, the listeners had no doubt that speech was presented

as a stimulus. Thus, these listeners presumably had knowledge, not only that the stimuli were speech, but of all the linguistic cues concomitant with speech. Such cues were probably not available to the listeners in the present study. Such a hypothesis suggests that linguistic cues are essential for an ear superiority to occur.

In addition to the above differences in experimental procedures between the various investigations, the response on the part of the listeners in previous research has required discrimination (speech or non-speech), recognition (which vowel, consonant, etc.), comprehension, and in some cases, retention (several stimulus presentations before responding), and recall. Recognition, comprehension, and retention certainly include linguistic information, which may be the necessary factor for the ear superiority to occur.

There seems little doubt that portions of the cortex are especially important for analysis of speech, language, linguistic, and psycholinguistic cues, etc. Analyses with the cortex, of course, have not been specified, but it seems clear that signals which are processed cortically (i.e., speech) and undergo refinements which are unavailable at subcortical regions could be expected to produce the ear superiority. In the present experiment, however, the stimuli are not treated as uniquely speech and probably possess no linguistic cues. The listeners' responses probably do not require recognition, comprehension, or retention, which are considered to be the results of cortical analysis. Last, but not least, the symmetrical anatomy and physiology of the subcortical auditory nervous system does not support ear superiority. The failure of this study to produce an ear superiority is most probably due to the fact that the stimuli used were simply acoustic events, which,

while being speech-like, undergo different neural processing than those experiments which demonstrate ear superiority effects.

Effects of Stimulus Types

Liberman, et al. (1967) referred to the class of phonemes that shows restructuring of their acoustic cues as a function of context as "encoded," and to those that show little or no restructuring as "unencoded." The same authors also suggested that the perception of these two classes is different. The results of the present study would, on the surface, tend to agree with such a dichotomy. The four different stimulus types used in this study (CV, V, VN, NBN) produced statistical differences in discrimination.

Of the two speech stimuli used in this study (CV and V), the CV stimulus can be considered an example of Liberman's "encoded" class. In the present study, the difference between the CV stimulus and the V stimulus was the 80 msec consonant transition at the beginning of the CV stimulus. It is, of course, the presence of the consonant transition that determines the "encoded" phoneme class. Although it cannot be determined, with any degree of certainty, what cues the listeners used to assign the titles "speech" and "non-speech" to the stimuli, it would seem logical that the 80 msec transition was important to this detection.

The CV stimulus was detected as speech a high percentage of the time regardless of the experimental condition. This is of particular interest in the dichotic presentations where the listeners generally reported speech whenever the CV stimulus appeared, whether in the directed ear, or not. This might be interpreted as concurring with Liberman's Speech Mode Hypothesis. However, a special speech mode is

not necessarily essential to this interpretation. An alternate hypothesis is that the auditory system processes the CV stimuli as simple acoustic events. Previous research has demonstrated that the auditory system has the capability to track pitch changes occurring in less than 80 msec that have no relation to speech stimuli (Sergeant and Harris, 1962; Harris, et al., 1963). Since the CV stimulus contains a transition at the beginning of the stimulus, which is nothing more than a complex signal exhibiting periodicity, with rapidly changing frequency contours, the auditory system may cue on this transition for speech or non-speech signals, when it is present.

A partial decoding of speech signals in the peripheral auditory system, by separation of periodicity and spectrum envelope information, has been suggested by several authors (Huggins, 1964; Stevens and House, 1968). Moreover, Stevens and House have suggested that listeners adopt a "set" or "mode" for speech perception which is governed almost entirely by linguistic segments and features which are apparently retained in the neural transformation of the acoustic signal. This would seem to be an important point in evaluating the results of the present research.

Obviously, there are several steps involved in the recognition of speech. Hirsh (1967) identifies four steps: detection, discrimination, recognition, and comprehension. In other words, speech must be detected before it can be discriminated. By the same token, various components of the speech signal must be discriminated prior to recognition, and recognition must precede comprehension. It would seem reasonable that the special speech mode would only pertain to recognition and comprehension, coupled with retention or short-term memory storage. In this case, the linguistic cues would be of utmost importance, and perhaps, result in

hemispheric dominance noted in previous research. At the detection level, as used in the present study, however, it would seem reasonable to suggest that the auditory system handles the signals as simple acoustic information, and differentiates on the basis of acoustic cues available in the signal. For example, the auditory system has available two kinds of information for peripheral processing. First, it has available the periodicity of the signal. The two speech-like stimuli exhibit periodicy, while the two noise stimuli are aperiodic. Second, the auditory system has available spectrum envelope information. In this case, only the CV stimulus exhibits a frequency and amplitude transition at the beginning of the stimulus. Thus, the auditory system could differentiate between the stimulus types on the basis of periodicity of the signal and the spectrum envelope information, without requiring additional linguistic information, and, thus, without requiring a special speech mode. The CV stimulus, with its consonantal transition, is detected as speech, and the two noise stimuli, with their aperiodic characteristics, are detected as non-speech.

The detection of the V stimulus lends support to this hypothesis. The V stimulus was periodic but had no transition. It was also confused as non-speech a significant percent of the time.

Confidence Ratings

The listeners' use of the confidence ratings and their subsequent ROC-curves and optimum criterion values tend to substantiate the above discussion. Although the listeners could perform the experimental task, they had little confidence in their responses. Only in the CV stimulus and, in one case, the NBN stimulus, did the listeners' confidence in

their responses vary significantly from chance. At the threshold of detectability, the listeners could differentiate between the stimulus types, possibly on the basis of periodicity and the information in the spectrum envelope. They were not confident, however, in assigning labels of "speech" and "non-speech" to the stimuli. The signals were apparently treated as simple acoustic events, without the assignment of the linguistic features which would contribute to greater confidence in the judgment of speech versus non-speech.

CHAPTER V

SUMMARY

In investigations using dichotic presentation of speech stimuli, recall of stimuli presented to the right ear has been significantly better than for the left ear in right-handed listeners. Previous investigators have always presented the speech stimuli at supra-threshold levels and required responses on the part of listeners that involved discrimination, recognition, comprehension, and, in some experiments, retention. A special speech mode has been suggested by some investigators to account for the retrieval of information in the speech signal that is encoded or restructured by the auditory nervous system. This study attempts to provide information on the relationship between monaural and binaural detection of speech, the ear superiority phenomenon, and the speech mode hypothesis, when speech stimuli are presented at the threshold of detectability.

Four different stimuli were used: a synthesized, consonant-vowel combination (CV), a synthesized, steady-state vowel (V), a vowel-shaped noise (VN), and a narrow-band noise (NBN). The stimulus types were randomized in such a way that six two-track tape recordings, with 100 stimuli on each track, were generated; two each, with a priori probabilities for the occurrence of speech of 0.25, 0.50, and 0.75. Stimuli were presented monaurally in the right and left ears; dichotically, with instructions to respond only on the basis of the stimuli in the

right ear; dichotically, with instructions to respond only on the basis of the stimuli in the left ear; and dichotically, with no specific instructions.

Ten listeners were tested individually, with the stimuli presented through earphones at a level of detectability determined for each ear immediately prior to the experimental condition. The listeners were instructed simply to listen for the presentation of a stimulus and to make a decision, based on their best guess, as to whether a stimulus was speech or non-speech. In addition, the listeners were instructed to make confidence judgments of their responses on a five-point scale.

Results of this study indicated that at the detection threshold listeners were able to differentiate between speech and non-speech stimuli with a high degree of accuracy for the CV, VN, and NBN stimuli. However, the V stimulus was detected both as speech and non-speech. No ear superiority was found for any of the a priori probabilities, stimulus types, or listening conditions.

Differences due to stimulus types were significant, with the V stimulus generally showing the lowest percent correct score. Analysis of the response errors indicated that the greatest number of errors were made when noise stimuli were in competition with a CV stimulus in the contralateral ear. The listeners appeared to track the CV stimulus, regardless of the ear to which it was presented or the instructions given the listeners. The confidence ratings indicated that the listeners could perform the experimental task although they had little confidence in their responses, except for the CV stimuli.

It was hypothesized that the results of the present study could be explained on the basis of periodicity and spectrum envelope information

in the acoustic signal, without the necessity of a special speech mode. In addition, it was hypothesized that the ear effect noted in previous research of dichotic presentation of speech stimuli might be due to an experimental paradigm which required recognition and comprehension on the part of the listeners which enabled them to use linguistic cues present in the stimuli.

APPENDICES

APPENDIX A

DESCRIPTION OF THE CV, V, VN, AND NBN STIMULI

Consonant-Vowel Combination

The CV stimuli were synthesized at Haskins Laboratory and consisted of an 80 msec consonant portion and a 220 msec vowel /a/ portion. The consonantal portion represented voiceless fricatives in the general vicinity between /f/ and /θ/. The hiss source varied in fundamental frequency from 116 Hz at time = 0, to 138 Hz at the beginning of the vowel portion. The relative overall amplitude of the consonantal portion varied from 28.0 dB to 3.0 dB at the beginning of the vowel portion.

The second and third formant frequency regions of the consonantal portion were variable in their initial frequency. The second formant consisted of eleven initial values from 543 Hz to 2837 Hz. The third formant consisted of eight initial values from 1190 Hz to 3530 Hz. Any particular CV stimulus could have any combination of initial values for F_2 and F_3, with F_1 constant for all CV stimuli. The order of occurrence for the different combinations was random throughout the test tapes.

The fundamental frequency of the vowel /a/ portion was 138 Hz. The first formant frequency for the vowel portion was 743 Hz, and the F_2 and F_3 regions were in the general area described by Peterson and Barney (1952).

Sixty-two synthesized CV stimuli were generated, each having a duration of exactly 300 msec and a 7 sec interstimulus interval. The onset of the first stimulus on the original CV tape was identified with the use of a magnetic tape reader. A distance, equivalent to 350 msec for a tape speed of 7½ in/sec, was measured in front of the initial stimulus, marked, and used as a starting point. The stimulus tape was reproduced from a tape recorder (Ampex, Model 601) to channel A of a second tape recorder (Ampex PR-10/2). The pulse gate output of an interval timer (Grason-Stadler, Model 471), adjusted for a 1 msec pulse with 7300 msec interstimulus interval, was fed to channel B of the second recorder. The original stimulus tape, positioned at the prearranged starting point, and the interval timer were started simultaneously. This procedure resulted in the CV stimuli being recorded on channel A of the second tape recorder, and a 1 msec pulse, occurring approximately 350 msec prior to the onset of every CV combination, being recorded on channel B. An oscilloscope (Hewlett-Packard, Model 122AR) was used to measure the time interval between the pulse and the onset of the CV stimuli. This time interval ranged from 325 msec to 375 msec, with the greatest majority being 350 msec.

The middle twenty pulse-stimulus combinations, which showed the smallest pulse to stimulus onset variations, were dubbed five times and spliced together, with the appropriate tape distance between them to maintain the 7 sec interstimulus interval. This procedure yielded a recording with 100 CV stimuli on channel A and their associated trigger pulses on channel B.

Vowel. - The steady-state vowel stimulus was synthesized on a passive vowel synthesizer at the Communication Sciences Laboratory

of the University of Florida. The synthesizer filters were adjusted for the regions of F_1, F_2, and F_3 of the vowel /a/, as described by Peterson and Barney (1952). The formant frequency regions were adjusted with the aid of an oscillator (Brüel and Kjaer, Model 1014) and a RMS vacuum tube voltmeter (Ballantine, Model 321). The frequency at the peak of each formant region was measured with a frequency counter (Hewlett-Packard, Model 522B). The peak formant frequency regions of the vowel were as follows: F_1 = 723 Hz, F_2 = 1082 Hz, and F_3 = 2455 Hz. The relative amplitudes of F_2 and F_3 were - 2.6 dB and - 20.0 dB, respectively. The formant frequency regions of the steady-state vowel /a/ can be seen in Figure 9.

Once the proper formant frequency regions and relative amplitudes for the vowel /a/ were achieved, the oscillator was replaced by a sawtooth pulse-train source with a frequency of 138 Hz (the same fundamental frequency of the vowel in the CV stimuli). The output of the synthesizer was recorded on a tape recorder (Ampex, Model 351).

Vowel-shaped noise. - The "voice" source for the synthesized steady-state vowel was replaced by a noise generator (Grasen-stadler, Model 901A), without readjusting the synthesizer settings. A shaped noise with the same energy concentrations and relative amplitudes as the steady-state vowel resulted. The output of the synthesizer was recorded on channel A of a tape recorder (Ampex, Model 602-2) and measured with the aid of a wave analyzer (General Radio, Model 1900-A). The wave analysis of the vowel-shaped noise can be seen in Figure 10.

Narrow-band noise. - The F_1 region of the vowel synthesizer, driven by the noise generator, was recorded on the second track of the tape recorder. This resulted in a shaped noise centered at 723 Hz,

but without F_2 and F_3 present. The wave analysis of this noise can be seen in Figure 11.

In summary, three master tape recordings were constructed:

1. CV stimuli on channel A and trigger pulses on channel B.

2. A continuous synthesized vowel /a/.

3. A continuous "vowel-shaped" thermal noise on channel A and a continuous "first formant" thermal noise on channel B.

These recordings were used to construct the actual stimulus recordings as described in Appendix B.

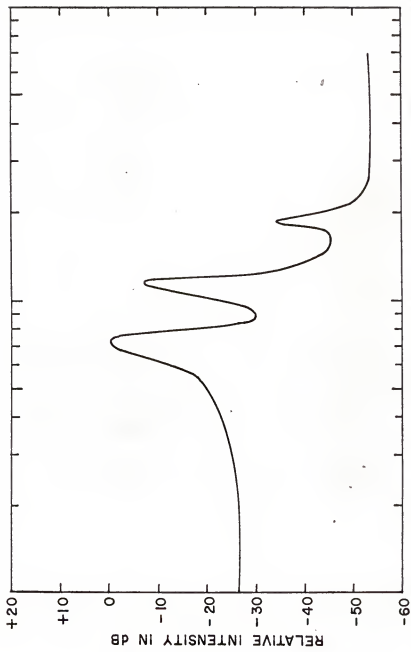

Figure 9. Formant frequency regions of the steady-state vowel stimulus /a/

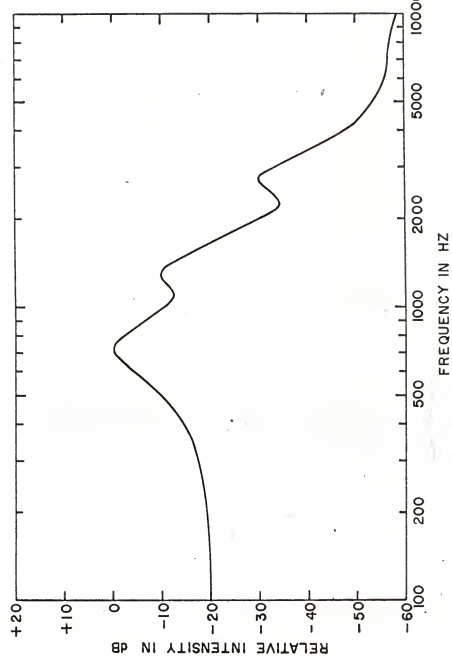

Figure 10. Formant frequency regions of the vowel-shaped noise stimulus VN

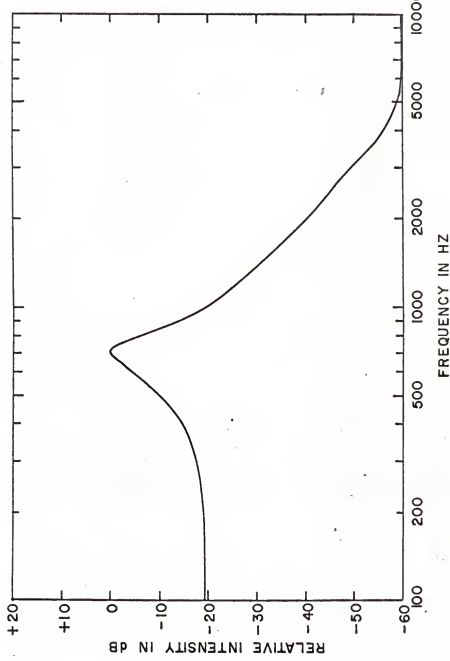

Figure 11. Formant frequency region of the narrow-band noise stimulus NBN

APPENDIX B

GENERATION OF TEST TAPES

Stimulus Distribution

The four stimulus types (CV, V, VN, and NBN) were randomized such that the a priori probability for the occurrence of a speech stimulus (CV or V) was 0.25, 0.50, or 0.75. Six tape recordings, two for each of the three a priori probabilities of speech occurrence, were developed as described later. In addition, the position of each competing stimulus in the dichotic test condition was randomized.

Tables 12 through 17 show the total number of each stimulus type on each channel of each test tape. In addition, for dichotic presentation, the total number of times the different stimuli are in competition with each other are shown. As an example, Table 12 shows that on channel A the CV combination occurs a total of 14 times; the V, 11 times; the VN, 38 times; and the NBN, 37 times. Combining CV and V, a speech-type stimulus is present 25 times (i.e., P(Sp = 0.25) on channel A of Tape I, while the two noise stimuli are present 75 times (i.e., P(Non-Sp) = 0.75). On channel B of Tape I, the CV is present 14 times; the V, 11 times; the VN, 35 times; and the NBN, 40 times. In addition, the table shows the CV is in competition with the CV on 4 occasions; the VN, 5 times; the NBN, 5 times, etc.

Table 18 shows the total number of stimulus comparisons. For example, a CV stimulus is in competition with another CV stimulus a total of 50 times and in competition with a V stimulus 39 times,

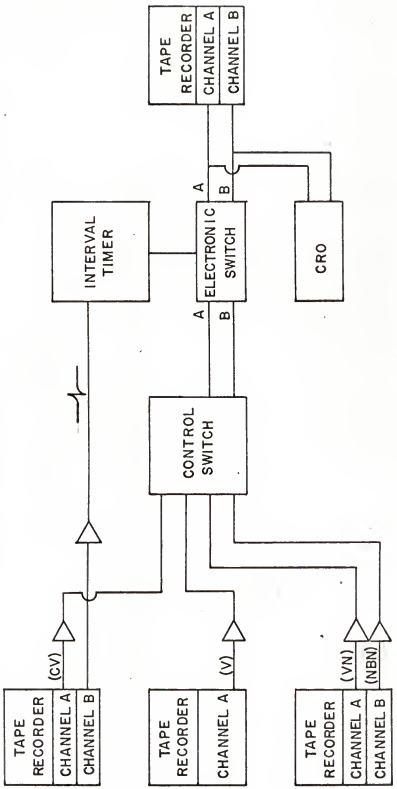

Figure 12. Schematic diagram of equipment used in generating the six test tapes

while a V stimulus from channel A is in competition with a CV stimulus from channel B 37 times and in competition with another V stimulus on channel B a total of 51 times. This total shows that a speech stimulus is in competition with another speech stimulus a total of 177 times. By the same token, a non-speech stimulus is in competition with another non-speech stimulus a total of 177 times, by summing all combinations of VN and NBN stimuli. Looking at all combinations of speech stimuli (i.e., CV and V) with non-speech stimuli (i.e., VN and NBN), it can be seen that a speech stimulus is in competition with a non-speech stimulus a total of 246 times.

Apparatus and Calibration

Figure 12 shows a schematic diagram of the equipment used in preparation of the test tapes. The CV stimuli (channel A) and the associated trigger pulses (channel B) were played from a tape recorder (Ampex, Model 354-2). The pulses were fed through an amplifier (McIntosh, Model MC 40) to the external trigger input of an interval timer (Grason-Stadler, Model 471) connected to an electronic switch (Grason-Stadler, Model 829 E). The trigger pulse from the CV tape opened the electronic switch for 300 msec (with a linear rise-fall time of 10 msec) after a fixed delay of 350 msec (i.e., the delay between the trigger pulse and the CV stimulus) and closed the switch for 7000 msec until the next trigger pulse. Thus it was possible, through suitable recording techniques described below, to place a stimulus on one tape channel simultaneously with any other stimulus on the second channel.

As reported in Appendix A, the three master tapes consisted of (1) CV and trigger pulses, (2) a continuously recorded synthesized vowel, (3) a continuously recorded vowel-shaped noise (VN) on one

channel, and (4) a continuously recorded narrow-band noise (NBN) on the second channel. The outputs of three tape recorders (see Figure 12) were led through amplifiers and a randomizing switch box to the two inputs of the electronic switch. The randomizing switch box consisted of several manually operated push-pull switches wired in such a way as to allow any combination of the four stimulus types to be passed to either input of the electronic switch. The outputs of the switch were tape recorded, and the six test tapes described above were generated.

The output voltage of each of the recorders was adjusted individually, with the gain controls on the amplifiers placed between the playback recorders and the control box. The output voltages were adjusted for equal peak-to-peak amplitude at the output of the electronic switch. The final tape recorder (Ampex, PR-10/2) was carefully calibrated.

In addition to the test stimuli, a 1000 Hz calibrating signal was recorded simultaneously on both channels of all recordings and used throughout the experiment to adjust the output voltage of the reproduce amplifiers in the Ampex PR-10/2 recorder. The voltage to the earphones, necessary to achieve a reference intensity level of 70 dB SPL for the 1000 Hz calibration signal, was measured with a condenser microphone (Brüel and Kjaer, Model 4132) and a 6 cc coupler. This voltage was monitored with a vacuum tube voltmeter (Ballantine, Model 321).

The acoustic output from the earphones was measured for overall amplitude differences between tapes and for amplitude differences between individual stimuli within any given test tape using a condenser microphone (Brüel and Kjaer, Model 4132), a 6 cc coupler, and a spectrometer (Brüel and Kjaer, Model 2112). The output of the spectrometer

was fed to a graphic level recorder (General Radio, Model 1521-B), and the peak amplitude of each stimulus on each tape was measured. The results of this measurement can be seen in Table 19. The test tapes varied in overall amplitude by ± 3.0 dB, and the standard deviation, reflecting variance within an individual tape, varied from 1.1 dB to 2.3 dB. The peak amplitude of the four stimulus types within a given test tape was within ± 1.4 dB of each other.

In addition to the intensity measurements, the frequency response characteristics of the earphones (Telephonics, TDH-39) and the frequency response characteristics of the tape-recorder-earphone combination were examined.

The differences in onset time, between individual pairs of stimuli on the test tapes, were measured to insure simultaneity of onset to the two ears in the dichotic condition. Two channels of a recorder (Sanborn, Model 150), with a paper speed of 100 mm/sec, were used for this measurement. This rapid speed allowed small differences (i.e., less than 5 msec) between onset times of channel A and channel B of the test tapes to be measured. Figure 13 shows the results of this measurement. It can be seen that 85 percent of the stimuli differ in onset times by less than 10 msec, with 58 percent having identical onset times.

A pilot study determined that the four stimulus types used in this experiment were equally detectable. The psychophysical method of limits was used and indicated that the threshold values of the four stimulus types were within ± 2.0 dB of each other.

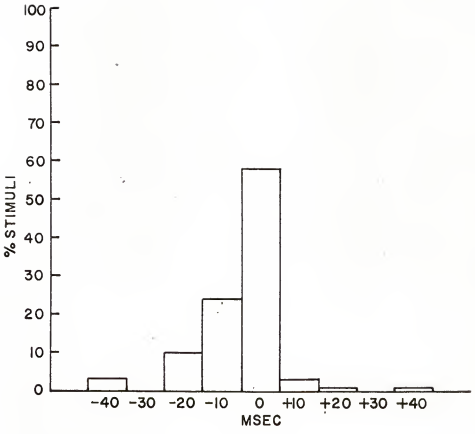

Figure 13. Distribution of onset time differences between channel A and channel B, in msec, measured with reference to channel A.

Table 12. Total number of the four stimulus types appearing on channel A and channel B of Test Tape I [P(Sp) = 0.25], and total number of times the four stimulus types are in competition with each other in the dichotic condition

Channel A	Channel B				
	CV	V	VN	NBN	Total Channel A
CV	4	0	5	5	14
V	0	3	4	4	11
VN	6	3	16	13	38
NBN	4	5	10	18	37
Total Channel B	14	11	35	40	100

Table 13. Total number of the four stimulus types appearing on channel A and channel B of Test Tape II [P(Sp) = 0.25], and total number of times the four stimulus types are in competition with each other in the dichotic condition

Channel A	Channel B				
	CV	V	VN	NBN	Total Channel A
CV	4	0	3	5	12
V	0	4	2	7	13
VN	4	4	12	18	38
NBN	5	4	14	14	37
Total Channel B	13	12	31	44	100

Table 14. Total number of the four stimulus types appearing on channel A and channel B of Test Tape III [P(Sp) = 0.50], and total number of times the four stimulus types are in competition with each other in the dichotic condition

Channel A	Channel B				
	CV	V	VN	NBN	Total Channel A
CV	4	6	7	8	25
V	7	7	9	2	25
VN	7	9	4	6	26
NBN	4	6	3	11	24
Total Channel B	22	28	23	27	100

Table 15. Total number of the four stimulus types appearing on channel A and channel B of Test Tape IV [P(Sp) = 0.50], and total number of times the four stimulus types are in competition with each other in the dichotic condition

Channel A	Channel B				
	CV	V	VN	NBN	Total Channel A
CV	10	4	6	4	24
V	7	4	10	5	26
VN	8	5	4	8	25
NBN	8	4	8	5	25
Total Channel B	33	17	28	22	100

Table 16. Total number of the four stimulus types appearing on channel A and channel B of Test Tape V [P(Sp) = 0.75], and total number of times the four stimulus types are in competition with each other in the dichotic condition

Channel A	Channel B				
	CV	V	VN	NBN	Total Channel A
CV	15	13	4	9	41
V	12	15	4	3	34
VN	4	7	3	0	14
NBN	6	3	1	1	11
Total Channel B	37	38	12	13	100

Table 17. Total number of the four stimulus types appearing on channel A and channel B of Test Tape VI [P(Sp) = 0.75], and total number of times the four stimulus types are in competition with each other in the dichotic condition

Channel A	Channel B				
	CV	V	VN	NBN	Total Channel A
CV	13	16	2	6	37
V	11	18	5	4	38
VN	4	4	2	1	11
NBN	4	5	3	2	14
Total Channel B	32	43	12	13	100

Table 18. Total number of the four stimulus types appearing on channel A and channel B of all Test Tapes combined, and total number of times the four stimulus types are in competition with each other in the dichotic condition

Channel A	Channel B				
	CV	V	VN	NBN	Total Channel A
CV	50	39	27	37	153
V	37	51	34	25	147
VN	33	32	41	46	152
NBN	31	27	39	51	148
Total Channel B	151	149	141	159	600

Table 19. Average relative intensity for channel A and channel B of each test tape

Test Tape	Channel	Average Relative Intensity In dB	SD
I	A	64.06	2.33
	B	66.00	1.43
II	A	62.17	1.55
	B	63.58	1.27
III	A	65.32	1.40
	B	62.73	1.53
IV	A	59.86	2.00
	B	62.40	1.96
V	A	61.67	1.42
	B	65.05	1.31
VI	A	64.81	1.39
	B	65.24	1.12

APPENDIX C

STATISTICAL SUMMARY TABLES

Table 20. Mean correct score and its standard deviation for the right and left ears for the monaural condition with [P(Sp) = 0.25]. The proportion scores represent arcsin transformations

Stimuli	Right Ear		Left Ear			
	Mean	S D	Mean	S D	Diff.	t*
CV	2.73	.70	2.93	.30	.20	1.26
V	1.72	1.13	1.83	.99	.11	.23
VN	2.12	.82	2.63	.41	.51	1.76
NBN	2.29	.92	2.83	.42	.54	1.69

*t = 2.88 at 0.01 confidence level.

Table 21. Mean correct score and its standard deviation for the right and left ears for the monaural condition with [P(Sp) = 0.50]. Scores represent arcsin transformations

Stimuli	Right Ear		Left Ear			
	Mean	S D	Mean	S D	Diff.	t*
CV	2.81	.51	2.82	.43	.01	.05
V	1.90	.98	1.84	.82	.06	.15
VN	2.45	.62	2.68	.35	.23	1.03
NBN	2.62	.51	2.82	.49	.20	.89

*t = 2.88 at 0.01 confidence level.

Table 22. Mean correct score and its standard deviation for right and left ears in the monaural condition with [P(Sp) = 0.75]. Scores represent arcsin transformations.

Stimuli	Right Ear		Left Ear			
	Mean	S D	Mean	S D	Diff.	t*
CV	2.64	.55	2.74	.38	.10	.48
V	1.49	1.27	1.75	1.02	.26	.51
VN	2.47	.96	2.14	1.00	.33	.75
NBN	2.38	.70	2.53	.75	.15	.46

*t = 2.88 at 0.01 confidence level.

Table 23. Summary table for the analysis of variance of the four stimulus types in the monaural conditions

Source of Variation	SS	df	MS	F
Stimulus Types	35.37	3	11.79	18.83**
Error	147.69	236	.626	
Total	183.06	239		

**$F_{.99}(3;236) = 3.78$.

Table 24. Summary table of Scheffé's test for a posteriori comparisons between means for monaural conditions

Stimulus Types		V	VN	NBN	CV
	Totals	105.42	144.97	154.82	166.80
V	105.42	--	39.55**	49.40**	61.38**
VN	144.97		--	9.85	21.83
NBN	154.82			--	11.98
CV	166.80				

**Critical value = $q_{.99}(3;236)\sqrt{n\ MS_{error}} = 29.24$.

Table 25. Summary table of the analysis of variance for the three *a priori* probabilities in the monaural conditions

Source of Variation	SS	df	MS	F
A Priori Prob.	2.29	2	1.14	1.50
Error	180.77	237	.76	
Total	183.06			

$F_{.99}(2;237) = 4.61$

$F_{.95}(2;237) = 3.00$

Table 26. Mean correct score and its standard deviation for binaural right and binaural left conditions with [P(Sp) = 0.25]. Scores represent arcsin transformations.

Stimuli	Right Ear		Left Ear			
	Mean	S D	Mean	S D	Diff.	t*
CV	2.98	.19	2.44	.72	.54	2.29
V	2.07	.76	1.62	1.14	.45	1.04
VN	2.15	.41	2.29	.65	.14	.58
NBN	2.50	.31	2.26	.55	.24	1.20

*t = 2.88 at 0.01 confidence level.

Table 27. Mean correct score and its standard deviation for binaural right and binaural left conditions with [P(Sp) = 0.50]. Scores represent arcsin transformations

Stimuli	Right Ear		Left Ear			
	Mean	S D	Mean	S D	Diff.	t*
CV	2.73	.44	2.36	.96	.36	1.11
V	1.89	.58	1.87	.81	.02	.06
VN	2.08	.41	2.15	.50	.07	.34
NBN	2.12	.46	2.16	.56	.04	.17

*t = 2.88 at 0.01 confidence level.

Table 28. Mean correct score and its standard deviation for binaural right and binaural left conditions with [P(Sp) = 0.75]. Scores represent arcsin transformations

Stimuli	Right Ear		Left Ear			
	Mean	S D	Mean	S D	Diff.	t*
CV	2.57	.45	2.72	.42	.15	.77
V	1.81	.52	1.83	.67	.02	.07
VN	1.78	.48	2.42	.58	.64	2.68
NBN	2.12	.54	1.84	.90	.28	.84

*t = 2.88 at 0.01 confidence level.

Table 29. Summary table of the analysis of variance of the four stimulus types in the binaural condition with directions

Source of Variation	SS	df	MS	F
Stimulus Types	18.94	3	6.31	15.39**
Error	97.83	236	.41	
Total	116.77	239		

**$F_{.99}(3;236) = 3.78$.

Table 30. Summary table of Scheffé's test for a posteriori comparisons between means in the binaural conditions with directions

Stimulus Types		V	VN	NBN	CV
	Totals	110.99	128.81	130.08	158.09
V	110.99	--	17.82	19.09	47.10**
VN	128.81		--	1.27	29.28**
NBN	130.08			--	28.01**
CV	158.09				--

**Critical value = $q_{.99}(3;236)\sqrt{n\,MS_{error}} = 23.66$

Table 31. Summary table of the analysis of variance of the three a priori probabilities in the binaural condition with directions

Source of Variation	SS	df	MS	F
A Priori Prob.	1.02	2	.51	1.04
Error	115.74	237	.49	
Total	116.76	239		

$F_{.99}(2;237) = 4.61.$

$F_{.95}(2;237) = 3.00.$

Table 32. Mean correct score and its standard deviations for binaural condition without directions with [P(Sp) = 0.50]. Scores represent arcsin transformations

Stimuli	Right Ear		Left Ear			
	Mean	S D	Mean	S D	Diff.	t*
CV	2.61	.43	2.64	.25	.03	.19
V	1.96	.61	1.82	.68	.14	.48
VN	1.72	.30	1.58	.60	.14	.66
NBN	2.10	.40	2.01	.25	.09	.60

*t = 2.88 at 0.01 confidence level.

Table 33. Mean correct score and its standard deviations for binaural condition without directions with [P(Sp) = 0.75]. Scores represent arcsin transformations

Stimuli	Right Ear		Left Ear			
	Mean	S D	Mean	S D	Diff.	t*
CV	2.60	.27	2.86	.24	.26	2.27
V	2.16	.65	2.03	.47	.13	.51
VN	1.56	.36	1.89	.52	.33	1.65
NBN	1.60	.59	1.31	.56	.29	1.13

*t = 2.88 at 0.01 confidence level.

Table 34. Mean correct score and its standard deviations for binaural condition without directions with [P(Sp) = 0.25]. Scores represent arcsin transformations

Stimuli	Right Ear		Left Ear			
	Mean	S D	Mean	S D	Diff.	t*
CV	2.68	.59	2.95	.27	.27	1.32
V	1.89	.79	1.45	.83	.44	1.21
VN	2.21	.17	2.25	.26	.04	.41
NBN	2.28	.11	2.33	.22	.05	.65

*t = 2.88 at 0.01 confidence level.

Table 35. Summary table of the analysis of variance of the four stimulus types in the binaural condition without directions

Source of Variation	SS	df	MS	F
Stimulus Types	31.01	3	10.34	34.47**
Error	71.36	236	.30	
Total	102.37	239		

**$F_{.99}(3;236) = 3.78$.

Table 36. Summary table of Scheffé's test for a posteriori comparisons between means in binaural condition without directions

Stimulus Type		VN	V	NBN	CV
	Totals	112.02	113.12	116.36	163.51
VN	112.02	--	1.10	4.34	51.49**
V	113.12		--	3.24	50.39**
NBN	116.36			--	47.15**
CV	163.51				--

**Critical value = $q_{.99}(3;236)\sqrt{n\ MS_{error}} = 20.22$.

Table 37. Summary table of analysis of variance for the three a priori probabilities in the binaural condition without directions

Source of Variation	SS	df	MS	F
A Priori Prob.	2.77	2	1.385	3.29*
Error	99.80	237	.421	
Total	102.57	239		

*$F_{.95}(2;237) = 3.00$.

Table 38. Summary table of Scheffé's test for a posteriori comparisons between means in the binaural condition without directions

A Priori Prob.		0.75	0.50	0.25
	Totals	160.19	164.57	180.25
0.75	160.19	--	4.38	20.06*
0.50	164.57		--	15.68
0.25	180.25			--

*Critical value = $q_{.95}(2;237)\sqrt{n\ MS_{error}} = 20.07$.

BIBLIOGRAPHY

1. Bartz, W.H., Satz, P., and Fennell, Eileen (1967a). "Grouping Strategies in Dichotic Listening: The Effects of Instructions, Rate and Ear Asymmetry," J. Exp. Psychol., 74, 132.

2. Bartz, W.H., Satz, P., Fennell, Eileen, and Lally, J.R. (1967b). "Meaningfulness and Laterality in Dichotic Listening," J. Exp. Psychol., 73, 204.

3. Broadbent, D.E. (1954). "The Role of Auditory Localization in Attention and Memory Span," J. Exp. Psychol., 47, 191.

4. Broadbent, D.E. (1957). "A Mechanical Model for Human Attention and Immediate Memory," Psychol. Rev., 64, 205-215.

5. Broadbent, D.E. (1958). Perception and Communication (Pergamon Press, London).

6. Bryden, M.P. (1962). "Order of Report in Dichotic Listening," Canad. J. Psychol., 16, 291.

7. Bryden, M.P. (1963). "Ear Preference in Auditory Perception," J. Exp. Psychol., 65, 103.

8. Bryden, M.P. (1966). "Short-Term Memory for Unbalanced Dichotic Lists," Psychol. Sci., 6, 379.

9. Carhart, R. (1967). "Binaural Reception of Meaningful Material," in Sensori-Neural Hearing Processes and Disorders, Henry Ford Hospt. International Symposium. A. Bruce Graham (Ed.) (Little, Brown and Co., Boston).

10. Carhart, R., Tillman, T., and Johnson, K.R. (1966). "Binaural Masking of Speech by Periodically Modulated Noise," J. Acoust. Soc. Amer., 39, 1037-1050.

11. Carhart, R., Tillman, T., and Johnson, K.R. (1967). "Release of Masking for Speech Through Interaural Time Delay," J. Acoust. Soc. Amer., 42, 124-138.

12. Cherry, E.C. (1953). "Some Experiments Upon the Recognition of Speech, With One and Two Ears," J. Acoust. Soc. Am., 25, 975-979.

13. Cherry, E.C., and Taylor, W.K. (1954). "Some Further Experiments Upon the Recognition of Speech, With One Ear and With Two Ears," J. Acoust. Soc. Am., 26, 554-559.

14. Clarke, Frank N. (1960). "Confidence Ratings, Second-Choice Responses and Confusion Matrices in Intelligibility Tests," J. Acoust. Soc. Am., 32, 35-46.

15. Delattre, P.D., Liberman, A.M., and Cooper, F.S. (1955). "Acoustic Loci and Transitional Cues for Consonants," J. Acoust. Soc. Am., 27, 769-773.

16. Dirks, D. (1964). "Perception of Dichotic and Monaural Verbal Material and Cerebral Dominance in Speech," Acta Oto-Laryngol., 58, 73.

17. Edwards, A.L. (1950). Experimental Design in Psychological Research (Rinehart, New York).

18. Egan, J.P. (1957). "Message Repetition, Operating Characteristics, and Confusion Matrices in Speech Communication," Technical Report AFCRC TR 57-50, AD 110064, Hearing and Communication Lab., Indiana Univ., June, 1957.

19. Eisenson, J. (1947). "Aphasics: Observations and Tentative Conclusions," J. Speech Dis., 12, 290-292.

20. Feldman, H. (1965). "Experiments on Binaural Hearing in Noise," Transl. Beltone Inst. No. 18, July, 1965.

21. Fex, J. (1967). "The Olivocochlear Feedback Systems," in Sensorineural Hearing Processes and Disorders, Henry Ford Hospt. International Symposium. A. Bruce Graham (Ed.) (Little, Brown and Co., Boston).

22. Gacek, R.R. (1967). "Afferent Auditory Neural System," in Sensorineural Hearing Processes and Disorders, Henry Ford Hospt. International Symposium. A. Bruce Graham (Ed.) (Little, Brown and Co., Boston).

23. Gerstman, G.L., and Kiang, N.Y.S. (1964). "Responses of Single Units in Auditory Cortex," Exptl. Neurol., 10, 1-18.

24. Hall, J.L., and Goldstein, M.H. (1968). "Representation of Binaural Stimuli by Single Units in Primary Auditory Cortex of Unanesthetized Cats," J. Acoust. Soc. Am., 43, 456-461.

25. Harris, J.D., Pikler, A.G., and Murphy, J.E. (1963). "Compensatory and Pursuit Tracking of Pitch," J. Acoust. Soc. Am., 35, 581-587.

26. Hirsh, I.J. (1967). "Perception of Speech," in Sensorineural Hearing Processes and Disorders, Henry Ford Hospt. International Symposium. A. Bruce Graham (Ed.) (Little, Brown and Co., Boston).

27. Huggins, A.W.F. (1964). "Distortion of the Temporal Pattern of Speech: Interruption and Alternation," J. Acoust. Soc. Am., 36, 1055-1064.

28. Kaiser, J.G., and David, E.E., Jr. (1960). "Reproducing the Cocktail Party Effect," J. Acoust. Soc. Am., 32, 918 (A).

29. Kimura, D. (1961a). "Some Effects of Temporal-Lobe Damage on Auditory Perception," Canad. J. Psychol., 15, 156.

30. Kimura, D. (1961b). "Cerebral Dominance and the Perception of Verbal Stimuli," Canad. J. Psychol., 15, 166.

31. Kimura, D. (1963a). "A Note on Cerebral Dominance in Hearing," Acta Oto-Laryngol., 56, 617.

32. Kimura, D. (1963b). "Speech Lateralization in Young Children as Determined by an Auditory Test," J. Comp. Physiol. Psychol., 56, 899.

33. Kimura, D. (1964). "Left-Right Differences in the Perception of Melodies," Quart. J. Exp. Psychol., 16, 355-358.

34. Kock, W.E. (1950). "Binaural Localization and Masking," J. Acoust. Soc. Am., 22, 801-804.

35. Levitt, H., and Rabiner, L.R. (1967a). "Binaural Release from Masking for Speech and Gain in Intelligibility," J. Acoust. Soc. Am., 42, 601.

36. Levitt, H., and Rabiner, L.R. (1967b). "Predicting Binaural Gain in Intelligibility and Release from Masking for Speech," J. Acoust. Soc. Am., 42, 820.

37. Liberman, A.M. (1957). "Some Results of Research on Speech Perception," J. Acoust. Soc. Am., 29, 117-123.

38. Liberman, A.M., Delattre, P.C., Cooper, F.S., and Gertsman, L.J. (1954). "The Role of Consonant-Vowel Transitions in the Perception of the Stop and Nasal Consonants," Psychol. Monographs, 68 (8, Whole No. 379).

39. Liberman, A.M., Cooper, F.S., Shankweiler, D.P., and Studdert-Kennedy, M. (1967). "Perception of the Speech Code," Psychol. Rev., 74, 431-461.

40. Miller, G.A., and Taylor, W.G. (1948). "The Perception of Repeated Bursts of Noise," J. Acoust. Soc. Am., 20, 171-182.

41. Milner, B. (1954). "Intellectual Function of the Temporal Lobe," Psychol. Bull., 51, 42-64.

42. Milner, B. (1958). "Psychological Defects Produced by Temporal Lobe Excision," Proc. Assoc. Res. Nerv. Ment. Dis., 36, 244-257.

43. Milner, B. (1962). "Laterality Effects in Audition," in Interhemispheric Relations and Cerebral Dominance. V.B. Mountcastle (Ed.) (John Hopkins Univ. Press, Baltimore).

44. Nielson, J.M. (1946). Agnosia, Apraxia, Aphasia (Paul Hoeber, New York, 2nd. Ed.).

45. Orr, D.B., Friedman, H.W., and Williams, J.C.C. (1965). "Trainability of Listening Comprehension of Speeded Discourse," J. Ed. Psychol., 56, 148-156.

46. Penfield, W., and Rasmussen, T. (1950). The Cerebral Cortex of Man (MacMillan Co, New York).

47. Penfield, W., and Roberts, L. (1959). Speech and Brain Mechanisms (Princeton Univ. Press, Princeton, N.J.).

48. Peterson, G.E., and Barney, H.L. (1952). "Control Methods Used in a Study of the Vowels," J. Acoust. Soc. Am., 23, 175-184.

49. Pollack, I., and Decker, L.R. (1958). "Confidence Ratings, Message Reception, and the Receiver Operating Characteristic," J. Acoust. Soc. Am., 30, 286-292.

50. Rasmussen, G.L. (1967). "Efferent Connections of the Cochlear Nucleus," in Sensorineural Hearing Processes and Disorders, Henry Ford Hospt. International Symposium. A. Bruce Graham (Ed.) (Little, Brown and Co., Boston).

51. Rosenzweig, M.R. (1951). "Representations of the Two Ears at the Auditory Cortex," Amer. J. Physiol., 67, 147-158.

52. Satz, P., Achenbach, K., Pattishall, E., and Fennell, E. (1965). "Order of Report, Ear Asymmetry and Handedness in Dichotic Listening," Cortex, 1, 377-396.

53. Schubert, E.D. (1956). "Some Preliminary Experiments on Binaural Time Delay and Intelligibility," J. Acoust. Soc. Am., 28, 895-901.

54. Schubert, E.D., and Schultz, M.C. (1962). "Some Aspects of Binaural Signal Selection," J. Acoust. Soc. Am., 34, 844-849.

55. Semmes, J., Weinstein, S., Ghent, L., and Teuber, H.L. (1960). Somatosensory Changes after Penetrating Brain Wounds in Man (Harvard Univ. Press, Cambridge).

56. Sergeant, R.L., and Harris, J.D. (1962). "Sensitivity to Unidirectional Frequency Modulation," J. Acoust. Soc. Am., 34, 1625-1628.

57. Shankweiler, D. (1966). "Effects of Temporal-Lobe Damage on Perception of Dichotically Presented Melodies," J. Comp. Physiol. Psychol., 62, 115-119.

58. Shankweiler, D. (1967). "Hemispheric Specialization and the Perception of Speech," Status Report on Speech Research, Haskins Laboratory, March, 1967.

59. Shankweiler, D., and Studdert-Kennedy, M. (1967). "Identification of Consonants and Vowels Presented to Left and Right Ears," Quart. J. Psychol., 19, 59-63.

60. Stevens, K.N., and House, A.S. (1968). "Speech Perception," in Foundations of Modern Auditory Theory. J. Tobias and E. Schubert (Eds.) (Academic Press, New York). In Press.

61. Stevens, K.N., Öhman, S.E.G., Studdert-Kennedy, M., and Liberman, A.M. (1964). "Cross-Linguistic Study of Vowel Discrimination," J. Acoust. Soc. Am., 36, 1989 (A).

62. Studdert-Kennedy, M., Liberman, A.M., and Stevens, K.N. (1963). "Reaction Time to Synthetic Stop Consonants and Vowels at Phoneme Centers and at Phoneme Boundaries," J. Acoust. Soc. Am., 35, 1900 (A).

63. Studdert-Kennedy, M., Liberman, A.M., and Stevens, K.N. (1964). "Reaction Time During the Discrimination of Synthetic Stop Consonants," J. Acoust. Soc. Am., 36, 1989 (A).

64. Tunturi, A.R. (1946). "A Study of the Pathway from the Medial Geniculate Body to the Acoustic Cortex in the Dog," Am. J. Physiol., 147, 311-319.

65. Walker, H.M., and Lev, J. (1953). Statistical Inference (Henry Holt and Co., New York).

66. Wallach, H., Newman, E.B., and Rosenzweig, M.R. (1949). "The Precedence Effect in Sound Localization," Am. J. Psychol., 62, 315-336.

67. Wepman, J. (1951). Recovery from Aphasia (Ronald Press, New York).

68. Weston, P.B., Miller, J.D., and Hirsh, I.J. (1965). "Release from Masking for Speech," J. Acoust. Soc. Am., 38, 1053-1054.

69. Winer, B.J. (1962). Statistical Principles in Experimental Design (McGraw-Hill Co., New York).

BIOGRAPHICAL SKETCH

William Grady Thomas was born March 21, 1934, at Charlotte, North Carolina. In June, 1952, he was graduated from West Mecklenburg High School. In August, 1957, he received the degree of Bachelor of Science with a major in Education of the Deaf from Appalachian State Teachers College. From 1957 until 1959, he served as a teacher at the North Carolina School for the Deaf. In 1959, he enrolled in the Graduate School of Washington University, St. Louis, Missouri. He held a traineeship from the Vocational Rehabilitation Administration until June, 1961, when he received the degree of Master of Arts with a major in Audiology. From 1961 until 1966, he served as Director of the Hearing and Speech Center and Assistant Professor of Audiology at the University of North Carolina School of Medicine. In 1966, he enrolled in the Graduate School of the University of Florida. He held an assistantship in the Department of Speech until January, 1968, and held a special research fellowship from the Office of Vocational Rehabilitation while pursuing his work toward the degree of Doctor of Philosophy.

William Grady Thomas is married to the former Joy Lynne Bell and is the father of three children. He is a member of the American Speech and Hearing Association, the Acoustical Society of America, the American Association for the Advancement of Science, and the National Rehabilitation Association.

This dissertation was prepared under the direction of the chairman of the candidate's supervisory committee and has been approved by all members of that committee. It was submitted to the Dean of the College of Arts and Sciences and to the Graduate Council, and was approved as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

August, 1968

Dean, College of Arts and Sciences

Dean, Graduate School

SUPERVISORY COMMITTEE:

Chairman